GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: September 19, 2002, 14:05:21 ; Search time 259.37 Seconds
(without alignments)
9609.226 Million cell updates/sec

Title: US-09-636-185-1
Perfect score: 1455
Sequence: 1 ggcaggcctaacacatgcaa ..... ctatagtgagtcgtattaca 1455

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1453	99.9	1455	22	AAF24493	Helper factor bact
2	1270.8	87.3	5105	20	AAX24989	E. coli MG1655 rrn
3	1259.6	86.6	1542	17	AAT29140	rRNA gene (rrsE) f
4	1259.6	86.6	1542	22	AAH75410	E. coli 16S rRNA.
5	1259.6	86.6	1542	22	AAC62270	Escherichia coli r
6	1258	86.5	1542	22	AAF23015	E. coli 16S rRNA s
7	1256.4	86.4	5097	20	AAX24983	E. coli MG1655 rrn
8	1256.4	86.4	5098	20	AAX24984	E. coli MG1655 rrn
9	1254	86.2	1541	22	AAS11022	Salmonella typhimu
10	1253.2	86.1	5014	20	AAX24987	E. coli MG1655 rrn
11	1250	85.9	1542	14	AAQ46119	E. coli 16S rRNA :
12	1249.2	85.9	1510	10	AAN01611	Escherichia coli 1
13	1247.6	85.7	1487	22	AAS11034	Shigella dysenteri
c 14	1245.2	85.6	5341	20	AAX24986	E. coli MG1655 rrn
15	1242.8	85.4	5018	20	AAX24985	E. coli MG1655 rrn
16	1238.2	85.1	1450	22	AAS11021	Escherichia coli 1
17	1237.4	85.0	1519	21	AAA65897	E. coli proliferat
18	1237.4	85.0	1549	21	AAA66050	E. coli proliferat
19	1237.4	85.0	1549	21	AAA66055	E. coli proliferat
20	1235.4	84.9	1542	17	AAT18759	E. coli 16S ribosc
c 21	1234.2	84.8	5090	20	AAX24988	E. coli MG1655 rrn
22	1225.1	84.2	1459	20	AAX82001	Pasteur sp. BCM 2
23	1223.1	84.1	1451	20	AAX82002	Pasteur sp. BCM 3
24	1200.8	82.5	1451	20	AAX81998	Pantoea sp BCM 1 p
25	1102	75.7	1474	21	AAI71794	Pseudoalteromonas
26	1099.4	74.9	1500	22	AAS11024	Vibrio cholera 16s
27	1054.4	72.5	1543	14	AAQ49109	16S rRNA gene of A
28	1049.6	72.1	1543	14	AAQ49110	16S rRNA gene of A
29	1048	72.0	1543	14	AAQ49111	16S rRNA gene of A
30	1044.6	71.8	1498	12	AAQ13085	16S rRNA gene cori
31	1040.8	71.5	5273	20	AAX24982	Haemophilus influe
c 32	1040.8	71.5	1830121	17	AAT42063	Haemophilus influe
33	1039.6	71.5	1487	22	AAS11033	Haemophilus in flu
34	1030.8	70.8	5519	20	AAX24981	Haemophilus influe
35	1028.4	70.7	1462	22	AAF85587	Pseudoalteromonas
36	1027.8	70.6	1545	22	AAF90012	Nucleotide sequenc
37	1020.4	70.1	1535	20	AAX83569	16S rDNA gene frag
38	1009.4	69.4	1532	22	AAF76234	Cycloclasticus pum
39	1006	69.1	1528	20	AAX83570	16S rDNA gene frag
40	1004	69.0	1518	17	AAT18764	Pseudomonas aerugi
41	999	68.7	1518	16	AAX84673	16 rDNA gene frag
42	997.1	68.5	1467	22	AAS11023	Pseudomonas aerugi
43	995.9	68.4	1528	22	AAF76235	Cycloclasticus pun
44	989.6	68.0	1501	22	AAI64907	Polycyclic aromati
45	989.6	68.0	1501	22	AAI64908	Polycyclic aromati

ALIGNMENTS

RESULT 1
AAF24493
ID AAF24493 standard; DNA; 1455 BP.
XX
AC AAF24493;
XX
DT 01-MAY-2001 (first entry)
XX
DE Helper factor bacterium DNA sequence.
XX
KW Helper factor bacterium; bacterial growth; Pasteuria; spore;
KW nematode control; crop; ds.
XX
OS Unidentified.
XX
PN WO200111017-A2.
XX
PD 15-FEB-2001.
XX
PF 10-AUG-2000; 2000WO-US21942.
XX
PR 10-AUG-1999; 99US-0148154.
XX
PA (ENTO-) ENTOMOS LLC.
XX
PI Gerber JF, White JH;
XX
DR WPI; 2001-191541/19.
XX
PT Producing bacterial Pasteuria endospores in vitro, useful in nematode
PT control programs, comprises growing bacteria in growth medium with
PT helper factor such as microorganism or chemical compound produced by a

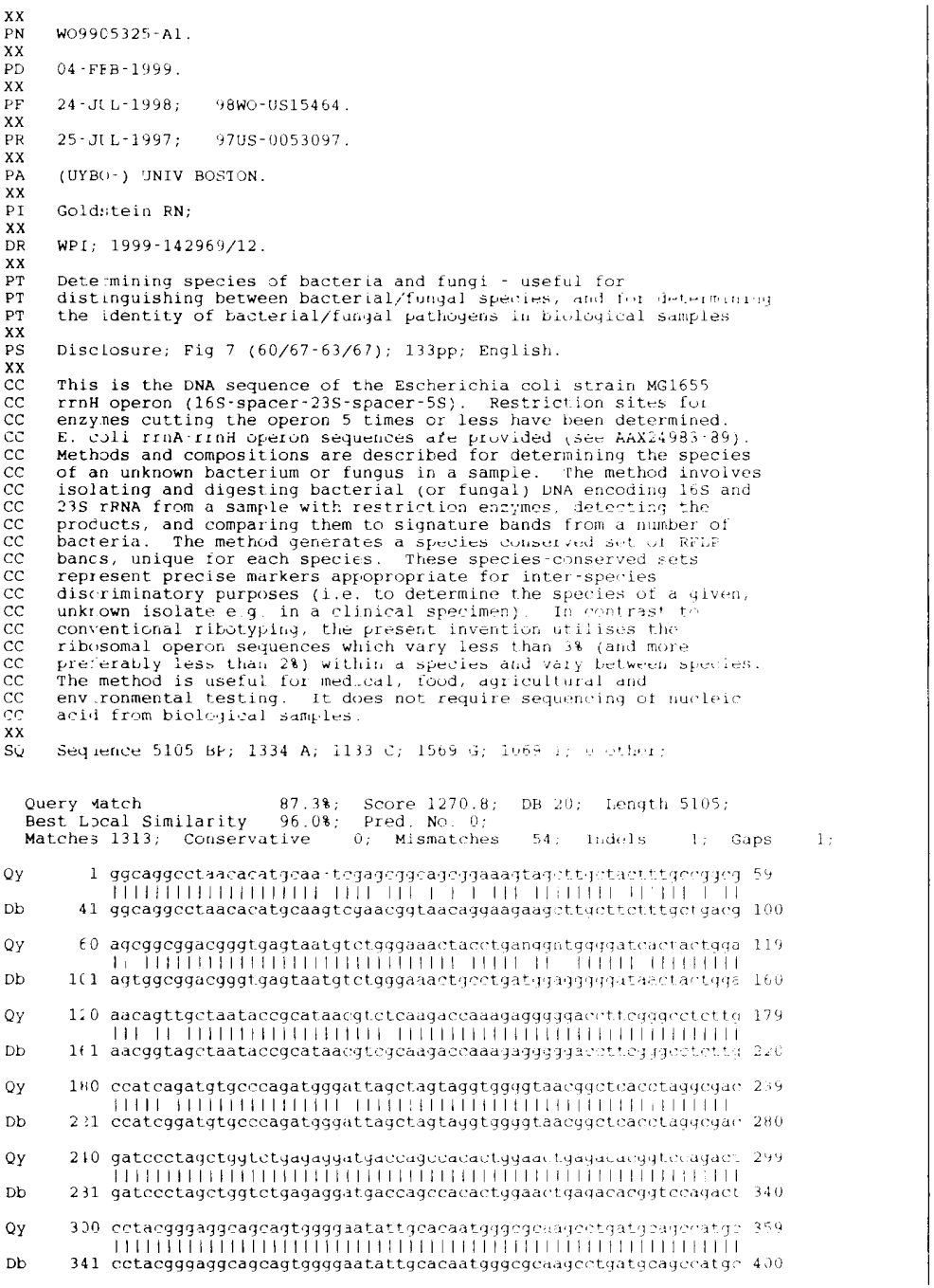
PT microorganism -
XX
PS Example 6; Page 23-24; 24pp; English.
XX
CC The present invention describes a method of producing Pasteuria spores in
CC vitro, involving growing the organism on growth media and obtaining the
CC endospores. These can then be applied to plants at risk of infection by
CC nematodes, where they grow and kill phytopathogenic nematodes which can
CC lead to crop losses. The present sequence comprises the coding sequence
CC of a bacterial helper factor capable of facilitating the in vitro growth
CC of Pasteuria.
XX
SQ Sequence 1455 BP; 365 A; 346 C; 454 G; 288 T; 2 other;

Query Match 99.9%; Score 1453; DB 22; Length 1455;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 2
AAX24989
ID AAX24989 standard; DNA; 5105 BP.
XX
AC AAX24989;
XX
DT 05-JUL-1999 (first entry)
XX
DE E. coli MG1655 rrnH operon (16S-spacer-23S-spacer-5S).
XX
KW Speciation; ribotyping; species discrimination; marker; RFLP;
KW restriction fragment length polymorphism; bacterium; fungus;
KW pathogen; rrnH operon; 16S RNA gene; 23S RNA gene; 5S RNA gene; ds.
XX
OS Escherichia coli.
XX
FH Key Location/Qualifiers
FT misc_feature 1..1542
FT /*tag= a
FT /label= 16S
FT misc_feature 1989..4892
FT /*tag= b
FT /label= 23S
FT misc_feature 4885..5105
FT /*tag= c
FT /label= 5S

XX
PN WO9905325-A1.
XX
PD 04-FEB-1999.
XX
PF 24-JUL-1998; 98WO-US15464.
XX
PR 25-JUL-1997; 97US-0053097.
XX
PA (UYBO-) UNIV BOSTON.
XX
PI Goldstein RN;
XX
DR WPI; 1999-142969/12.
XX
PT Determining species of bacteria and fungi - useful for
PT distinguishing between bacterial/fungal species, and for determining
PT the identity of bacterial/fungal pathogens in biological samples
XX
PS Disclosure; Fig 7 (60/67-63/67); 133pp; English.
XX
CC This is the DNA sequence of the Escherichia coli strain MG1655
CC rrnH operon (16S-spacer-23S-spacer-5S). Restriction sites for
CC enzymes cutting the operon 5 times or less have been determined.
CC E. coli rrnA-rrnH operon sequences are provided (see AAX24983-89).
CC Methods and compositions are described for determining the species
CC of an unknown bacterium or fungus in a sample. The method involves
CC isolating and digesting bacterial (or fungal) DNA encoding 16S and
CC 23S rRNA from a sample with restriction enzymes, detecting the
CC products, and comparing them to signature bands from a number of
CC bacteria. The method generates a species conserved set of RFLP
CC bands, unique for each species. These species-conserved sets
CC represent precise markers appropriate for inter-species
CC discriminatory purposes (i.e. to determine the species of a given,
CC unknown isolate e.g. in a clinical specimen). In contrast to
CC conventional ribotyping, the present invention utilises the
CC ribosomal operon sequences which vary less than 3% (and more
CC preferably less than 2%) within a species and vary between species.
CC The method is useful for medical, food, agricultural and
CC environmental testing. It does not require sequencing of nucleic
CC acid from biological samples.
XX
SQ Sequence 5105 BP; 1334 A; 1133 C; 1569 G; 1069 T; 0 other;

Query Match 87.3%; Score 1270.8; DB 20; Length 5105;
Best Local Similarity 96.0%; Pred. No. 0;
Matches 1313; Conservative 0; Mismatches 54; Indels 1; Gaps 1;



RESULT 3
AAT29140
ID AAT29140 standard; DNA; 1542 BP.
XX

```
AC   AAT29140;
XX
DT   02-DEC-1996  (first entry)
XX
DE   rRNA gene (rrsE) from Escherichia coli.
XX
KW   p53; mutant; mutation; cleavage, nuclease, cleavase, Thermus;
KW   Escherichia; Saccharomyces; Campylobacter; Mycobacterium; Shigella;
KW   Staphylococcus; identification; detection; ds.
XX
OS   Escherichia coli.
XX
PN   WO9605267-A1.
XX
PD   23-MAY-1996.
XX
PF   09-NOV-1995;   95WO-US14673.
XX
PR   30-AUG-1995;   95US-0520946.
PR   09-NOV-1994;   94US-0337164.
PR   09-MAR-1995;   95US-0402601.
PR   07-JUN-1995;   95US-0484956.
XX
PA   (THIR-) THIRD WAVE TECHNOLOGIES INC.
XX
PI   Brow MAD, Dahlberg JE, Fors L, Heisler LM, Lyamichev VI;
PI   Oldenburg MC, Olive DM;
XX
DR   WPI; 1996-259862/26.
XX
PT   Cleavage of nucleic acids to detect mutation(s) - allows detection
PT   esp. in human p53 gene, to identify strains of microorganisms and
PT   viruses
XX
PS   Example 34, Page 314; 433pp; English.
XX
CC   Cleavage of nucleic acids using an enzyme, especially a nuclease
CC   selected from the group consisting of Cleavase (RTM) BN enzyme,
CC   Thermus aquaticus DNA polymerase, Thermus thermophilus DNA
CC   polymerase, Escherichia coli ExoIII and the Saccharomyces cerevisiae
CC   Rad1/Rad10 complex.  The nucleic acid substrate is preferably an
CC   oligonucleotide containing a human p53 gene sequence or
CC   alternatively, microbial gene sequences.  Cleavage products are
CC   compared to the cleavage products of reference gene sequences.  The
CC   method is used for detecting mutation in the human p53 gene; for
CC   identifying strains of microorganisms, especially bacteria selected
CC   from the the group of members of the genera Campylobacter,
CC   Escherichia, Mycobacterium, Salmonella, Shigella and Staphylococcus.
CC   The method may also be used for the identification of viruses,
CC   especially hepatitis C virus (HCV) and simian immunodeficiency virus
CC   (SIV).  Seven primers (AAT29133-39) were used to amplify 16S rRNA
CC   genes.  The amplified genes were then subjected to the cleavage method
CC   and the resulting fragmentation patterns used in identification of
CC   the microorganisms from which the rRNA genes originated.  16S rRNA
CC   genes are given in AAT29140-42
XX
SQ   Sequence 1542 BP; 389 A; 352 C; 487 G; 314 T; 0 other.
```

```
  Query Match             86.6%;  Score 1259.6;  DB 17;  Length 1542;
  Best Local Similarity   95.5%;  Pred. No. 0,
  Matches 1306;  Conservative    0;  Mismatches   61;  Indels    1;  Gaps    1;

Qy       1 ggcaggcctaacacatgcaa tcgagcggcagcggaaaagtagcttgctactttgccggcg 59
           |||||||||||||||||||| |||| ||| | | | ||| |||||||| |||||| | ||
Db       4 ggcaggcctaacacatgcaagtcgaacggtaacaggaagaagcttgcttctttgctgacg 100

Qy      60 agcggcggacgggtgagtaatgtctgggaaactgcctgatggagggggataactactgga 119
           || |||||||||||||||||||||||||||||||| ||||| ||  |||||| ||||||||
Db     10 agtggcggacgggtgagtaatgtctgggaaactgcctgatggagggggataactactgga 160

Qy     120 aacagttgctaataccgcataacgtctcaagaccaaagagggggaccttcgggcctcttg 179
           ||| || |||||||||||||||| || ||||||||||||||||||||||||||||||||
```

```
Db     161 aacggtagcttaatacgcatgatcgtgcgagagggaaagatttatcgccgatagatgagc 220

Qy     180 ctatcagattagctagttggtgaggtaatggctcaccaaggcgacgatccgtaactggtct 239
           |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     221 ctatcggattgtgcttcaagtgggattagctagtaggtgaggtaacggctcacctaggcga 280

Qy     240 cgatccgtaactggtctgagaggatgatcagccacactggaactgagacacggtccagact 299
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     281 gatccctagctggtctgagaggatgaccagccacactggaactgagacacggtccagact 340

Qy     300 cctacgggaggcagcagtgaggaatattggacaatgggcgcaagcctgatccagccatgcc 359
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     341 cctacgggaggcagcagtggggaatattgcacaatgggcgcaagcctgatgcagccatgcc 400

Qy     360 cgcgtgtatgaagaaggccttcgggttgtaaagtactttcagcggggaggaagggagtaaa 419
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     401 cgcgtgtatgaagaaggccttcgggttgtaaagtactttcagcggggaggaagggagtaaa 460

Qy     420 gttaatacctttgctcattgacgttacccgcagaagaagcaccggctaactccgtgccag 479
           |||||| | |    |||||||||||||||||| |||||||| |||||| | | | | |
Db     461 agttaatacctttgctcattgacgttacccgcagaagaagcaccggctaactctgtgccag 520

Qy     480 cagccgccgcggtaatacagaggggtgcaagcgttaatcggaattactgggcgtaaagcgc 539
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     521 cagccgccgcggtaatacggaggggtgcaagcgttaatcggaattactgggcgtaaagcgc 580

Qy     540 gcaggcggttctttaagtctgatgtgaaagcccccgggctcaacctgggaactgcattgga 599
           |||||||| ||| ||||| ||||||||||||||||||||||||||||||||| |||||| ||
Db     581 gcaggcggtttgttaagtcagatgtgaaatccccgggctcaacctgggaactgcatctga 640

Qy     600 aactggcaggctagagtcttgtagaggggggtagaattccaggtgtagcggtgaaatgcgt 659
           |||||| ||| |||||| ||||||||||||||||||||||||||||||||||||||||||
Db     641 tactggcaagcttgagtctcgtagaggggggtagaattccaggtgtagcggtgaaatgcgt 700

Qy     660 agagatctggaggaataccggtggcgaaggcggccccctggacgaagactgacgctcagat 719
           |||||||||||| |||||||||||||||||||||||||||||||||||||||||||||||
Db     701 agagatctggaggaataccggtggcgaaggcggccccctggacgaagactgacgctcagat 760

Qy     720 gtgcgaaagcgtggggagcaaacaggattagataccctggtagtccacgccgtaaacgatg 779
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     761 gtgcgaaagcgtggggagcaacaggattagataccctggtagtccacgccgtaaacgatg 820

Qy     780 gtcgacttggaggttgtgcccttgaggcgtggcttccggagctaacgcgttaagtcgacc 839
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     821 gtcgacttggaggttgtgcccttgaggcgtggcttccggagctaacgcgttaagtcgacc 880

Qy     840 gcctggggagtacggccgcaaggttaaaactcaaatgaattgacgggggcccgcacaagcg 899
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     881 gcctggggagtacggccgcaaggttaaaactcaaatgaattgacgggggcccgcacaagcg 940

Qy     900 ggtggagcatgtggtttaattcgatgcaacgcgaagaaccttacctggtcttgacatccac 959
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     941 ggtggagcatgtggtttaattcgatgcaacgcgaagaaccttacctggtcttgacatccac 1000

Qy     960 agaactttccagagatggattggtgccttcgggaactgtgagacaggtgctgcatggctgt 1019
           ||| ||| |||||||| |||||||||||||||| |||||||||||||||||||||||||||
Db    1001 agaactttccagagatgaattggtgccttcgggaactgtgagacaggtgctgcatggctgt 1060

Qy    1020 cgtcagctcgtgttgtgaaatgttgggttaagtcccgcaacgagcgcaacccttatcctt 1079
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1061 cgtcagctcgtgttgtgaaatgttgggttaagtcccgcaacgagcgcaacccttatcctt 1120

Qy    1080 tttgttgccagcggttaggccgggaactcaaaggagactgccagtgataaactggaggaag 1139
           |||||||||||||| |||||||||||||||||||||||||||||||| |||||||||||||
Db    1121 tttgttgccagcggtccggccgggaactcaaaggagactgccagtgataaactggaggaag 1180

Qy    1140 gtggggatgacgtcaagtcatcatggcccttacgaccagggctacacacgtgctacaatgg 1199
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1181 gtggggatgacgtcaagtcatcatggcccttacgaccagggctacacacgtgctacaatgg 1240

Qy    1200 cgcatacaaagagaagcgacctcgcgagagcaagcggacctcataaagtgcgtcgtagtcc 1259
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1241 cgcatacaaagagaagcgacctcgcgagagcaagcggacctcataaagtgcgtcgtagtcc 1300
```

```
Qy    1260 tccggattggagtctgcaactcgactccatgaagtcggaatcgctagtaatcgtagatca 1319
           ||||||||||||||||||||||| |||||||||||||||||||||||||||| | |||||
Db     130 tccggattggagtctgcaactcgactccatgaagtcggaatcgctagtaatcgtggatca 1360

Qy    1320 gaatgctacggtgaatacgttcccgggccttgtacacaccgcccgtcaa 1367
           |||||| |||||||||||||||| ||||||||||||||||||||||| |
Db     136 gaatgccacggtgaatacgttcccgggccttgtacacaccgcccgtca 1408


RESULT   4
AAH75410
ID   AAH75410 standard; rRNA; 1542 BP.
XX
AC   AAH75410;
XX
DT   18-OCT-2001  (first entry)
XX
DE   E. coli 16S rRNA.
XX
KW   16S rRNA; 23S rRNA; RNA binding; antimicrobial; ss.
XX
OS   Escherichia coli.
XX
FH   Key             Location/Qualifiers
FT   misc_binding    9..13
FT                   /*tag= a
FT                   /bound_moiety= "16S rRNA"
FT                   /note= "Binds nucleotides 25-21 to form a duplex"
FT   misc_binding    17..19
FT                   /*tag= b
FT                   /bound_moiety= "16S rRNA"
FT                   /note= "Binds nucleotides 918-916 to form a duplex"
FT   misc_binding    21..25
FT                   /*tag= c
FT                   /bound_moiety= "16S rRNA"
FT                   /note= "Binds nucleotides 13-9 to form a duplex"
FT   misc_binding    27..37
FT                   /*tag= d
FT                   /bound_moiety= "16S rRNA"
FT                   /note= "Binds nucleotides 556-547 to form a duplex"
FT   misc_binding    39..46
FT                   /*tag= e
FT                   /bound_moiety= "16S rRNA"
FT                   /note= "Binds nucleotides 403-394 to form a duplex"
FT   misc_binding    52..58
FT                   /*tag= f
FT                   /bound_moiety= "16S rRNA"
FT                   /note= "Binds nucleotides 359-354 to form a duplex"
FT   stem_loop       61..106
FT                   /*tag= g
FT   misc_binding    113..115
FT                   /*tag= h
FT                   /bound_moiety= "16S rRNA"
FT                   /note= "Binds nucleotides 814-812 to form a duplex"
FT   misc_binding    123..142
FT                   /*tag= i
FT                   /bound_moiety= "16S rRNA"
FT                   /note= "Binds nucleotides 238-221 to form a duplex"
FT   stem_loop       144..178
FT                   /*tag= j
FT   stem_loop       184..193
FT                   /*tag= k
FT   stem_loop       199..218
FT                   /*tag= l
FT   misc_binding    221..238
FT                   /*tag= m
FT                   /bound_moiety= "16S rRNA"
FT                   /note= "Binds nucleotides 142-123 to form a duplex"
FT   stem_loop       240..286
FT                   /*tag= n
FT   stem_loop       289..311
FT                   /*tag= o
FT   misc_binding    312..314
FT                   /*tag= p
FT                   /bound_moiety= "16S rRNA"
FT                   /note= "Binds nucleotides 115-113 to form a duplex"
FT   stem_loop       316..337
FT                   /*tag= q
FT   stem_loop       339..350
FT                   /*tag= r
FT   misc_binding    354..359
FT                   /*tag= s
FT                   /bound_moiety= "16S rRNA"
FT                   /note= "Binds nucleotides 58-52 to form a duplex"
FT   stem_loop       367..392
FT                   /*tag= t
FT   misc_binding    394..405
FT                   /*tag= u
FT                   /bound_moiety= "16S rRNA"
FT                   /note= "Binds nucleotides 46-49 to form a duplex"
FT   stem_loop       406..436
FT                   /*tag= v
FT   stem_loop       442..492
FT                   /*tag= w
FT   stem_loop       500..545
FT                   /*tag= x
FT   protein_bind    523
FT                   /*tag= y
FT                   /bound_moiety= "Streptomycin"
FT   misc_binding    547..556
FT                   /*tag= z
FT                   /bound_moiety= "16S rRNA"
FT                   /note= "Binds nucleotides 37-27 to form a duplex"
FT   misc_binding    567..569
FT                   /*tag= aa
FT                   /bound_moiety= "16S rRNA"
FT                   /note= "Binds nucleotides 883-881 to form a duplex"
FT   misc_binding    577..586
FT                   /*tag= ab
FT                   /bound_moiety= "16S rRNA"
FT                   /note= "Binds nucleotides 764-755 to form a duplex"
FT   stem_loop       588..651
FT                   /*tag= ac
FT   misc_binding    655..671
FT                   /*tag= ad
FT                   /bound_moiety= "16S rRNA"
FT                   /note= "Binds nucleotides 751-735 to form a duplex"
FT   stem_loop       672..717
FT                   /*tag= ae
FT   protein_bind    693
FT                   /*tag= af
FT                   /bound_moiety= "Pactamycin, Edeine"
FT   stem_loop       725..732
FT                   /*tag= ag
FT   misc_binding    735..751
FT                   /*tag= ah
FT                   /bound_moiety= "16S rRNA"
FT                   /note= "Binds nucleotides 671-655 to form a duplex"
FT   misc_binding    755..764
FT                   /*tag= ai
FT                   /bound_moiety= "16S rRNA"
FT                   /note= "Binds nucleotides 586-577 to form a duplex"
FT   stem_loop       769..810
FT                   /*tag= aj
FT   protein_bind    794
FT                   /*tag= ak
FT                   /bound_moiety= "Kasugamycin, Pactamycin, Edeine"
FT   protein_bind    795
FT                   /*tag= al
FT                   /bound_moiety= "Kasugamycin, Pactamycin, Edeine"
FT   misc_binding    821..826
FT                   /*tag= am
FT                   /bound_moiety= "16S rRNA"
FT                   /note= "Binds nucleotides 979-974 to form a duplex"
FT   stem_loop       829..857
```

```
FT                   /*tag=  an
FT   stem_loop       861..868
FT                   /*tag=  ao
FT   protein_bind    864
FT                   /*tag=  ap
FT                   /bound_moiety= "Oxazolidinone"
FT   misc_binding    874..879
FT                   /*tag=  aq
FT                   /bound_moiety= "16S rRNA"
FT                   /note= "Binds nucleotides 826-82[illegible] to form a duplex"
FT   misc_binding    881..883
FT                   /*tag=  ar
FT                   /bound_moiety= "16S rRNA"
FT                   /note= "Binds nucleotides 569-56[illegible] to form a duplex"
FT   stem_loop       885..912
FT                   /*tag=  as
FT   protein_bind    910
FT                   /*tag=  at
FT                   /bound_moiety= "Streptomycin"
FT   protein_bind    911
FT                   /*tag=  au
FT                   /bound_moiety= "Streptomycin"
FT   protein_bind    913
FT                   /*tag=  av
FT                   /bound_moiety= "Streptomycin"
FT   protein_bind    914
FT                   /*tag=  aw
FT                   /bound_moiety= "Streptomycin"
FT   protein_bind    915
FT                   /*tag=  ax
FT                   /bound_moiety= "Streptomycin"
FT   misc_binding    916..918
FT                   /*tag=  ay
FT                   /bound_moiety= "16S rRNA"
FT                   /note= "Binds nucleotides 19-17 to form a duplex"
FT   misc_binding    921..933
FT                   /*tag=  az
FT                   /bound_moiety= "16S rRNA"
FT                   /note= "Binds nucleotides 1396-1384 to form a duplex"
FT   protein_bind    926
FT                   /*tag=  ba
FT                   /bound_moiety= "Kasugamycin, Edeine"
FT   misc_binding    938..943
FT                   /*tag=  bb
FT                   /bound_moiety= "16S rRNA"
FT                   /note= "Binds nucleotides 1345-1340 to form a duplex"
FT   misc_binding    946..955
FT                   /*tag=  bc
FT                   /bound_moiety= "16S rRNA"
FT                   /note= "Binds nucleotides 1235-1225 to form a duplex"
FT   stem_loop       960..975
FT                   /*tag=  bd
FT   misc_binding    984..990
FT                   /*tag=  be
FT                   /bound_moiety= "16S rRNA"
FT                   /note= "Binds nucleotides 1221-1215 to form a duplex"
FT   misc_binding    997..1003
FT                   /*tag=  bf
FT                   /bound_moiety= "16S rRNA"
FT                   /note= "Binds nucleotides 1044-1037 to form a duplex"
FT   stem_loop       1007..1022
FT                   /*tag=  bg
FT   stem_loop       1025..1036
FT                   /*tag=  bh
FT   misc_binding    1037..1044
FT                   /*tag=  bi
```

```
  Query Match             86.6%;  Score 1259.6;  DB 22;  Length 1542;
  Best local Similarity   76.8%;  Pred. No. 0;
  Matches 1050;  Conservative  256;  Mismatches   61;  Indels    1;  Gaps    1;

Qy       1 ggcaggcctaacacatgcaa-tcgaacggcaacgaaaagtagcttgctacttgccggcg 59
           |||||||| |||| |:|||| :||| ||| | | | ||| |||::|:: |:::|| | ||

Db      41 ggcaggccttacacatgcaagtcgaacggtaacaggaagaagcttgcttctttgctgacgag 100

Qy      60 agcggcggacgggtgagtaatgtctgggaaactgcctgatggagggggataactactggaa 119
           || |||||||||:|:||:|:|:|||||||: |:|| || |||||:|:|:||||

Db     101 agtggcggacgggtgagtaatgtctgggaaactgccctgatggagggggataactactgga 160

Qy     120 aacggttgctaataccgcataacgtcttcggaccaaagagggggaccttcgggcctcttgcc 179
           ||| |: ||:||:|||||:||||:| ||||||||||:|||||||| ::| ||| ::::

Db     161 aacggtagctaataccgcataacgtcgcaagaccaaagagggggaccttcgggcctcttgcc 220

Qy     180 atcagatgtgcccagatgggattagctagtaggtggggtaacggctcacctaggcgac 239
           ||||| ||:|:||||:|||:||||::|||:|||:|.||:|||||||:|||:|:|||

Db     221 atcagatgagcccagaagggattagctagtaggtggggtaacggctcacctaggcgac 280

Qy     240 gatcc[illegible] 299
           ||:|||:|||:||:|||||||:||||||||||:|||||:|||| |||:|||||:

Db     281 gatccctagctggtctgagaggatgaccagccacactggaactgagacacggtccagact 340

Qy     300 cctacgggaggcagcagtggggaatattgcacaatgggcgcaagcctgatgcagccatgc 359
           ||:|||||||||||||:|||||:|::||||||:|||||||||| ||:|||:|||||

Db     341 cctacgggaggcagcagtggggaatattgcacaatgggcgcaagcctgatgcagccatgc 400

Qy     360 cgcgtgtatgaagaaggccttcgggttgtaaagtactttcagcggggaggaagggagtaaa 419
           ||||:|:|:||||||||||:||||:|:||||:|:|||||||:|||||| : |

Db     401 cgcgtgaatgaagaagccttcgggttgtaaagtactttcagcggggaggaagggagtaaa 460

Qy     420 ggttaataacctt[illegible] 479
           |::|||:| | :    |: ||||:||||:|||||||||||||||:|||:||:|:|

Db     461 agttaataccttggtcattgacgttacccgcagaagaagcaccggctaactccgtgccag 520

Qy     480 cagccgcggtaatacggagggtgcaagcgttaatcggaattactgggcgtaaagcgcac 539
           ||||||||||:||:||||||||:|||||||::|:||||::||:||||:|||||||

Db     521 cagccgcggtaatacggagggtgcaagcgttaatcggaattactgggcgtaaagcgcac 580

Qy     540 gcaggcggt[illegible] 599
           |||||||| :|: |||| ||:|:||||:|||||||:|||||:|||| :||| |

Db     581 gcaggcggtttgttaagtcagatgtgaaagccccgggctcaacctgggaactgcatctga 640

Qy     600 aactggcaggctagagtcttgtagaggggggtagaattccaggtgtagcggtgaaatgcg 659
           ||:|||| || ||:|||: |:|||||||||:||||::|||||:||| |:|:|:||.

Db     641 aactggcaagctggagtctcgtagaggggggtagaattccaggtgtagcggtgaaatgcg 700

Qy     660 tagagatctggaggaataccggtggcgaaggcggccccctgaacagattgacgctgagg 719
           :|||||:|:|||||||:||||:|||||||||||:||||:|||| |||||:|||:|

Db     701 tagagatcaggaggaacaccagtggcgaaggcgactctctggtctgtaactgacgctgag 760

Qy     720 gtgcgaaagcgtggggagcaaacaggattagataccctggtagtccacgccgtaaacgat 779
           |:|||||||:|||||||||||||||:|||:|||||:| :|||||||:|||||||:|

Db     761 gtgcgaaagcgtggggagcaaacaggattagataccctggtagtccacgccctaaacgat 820

Qy     780 gtcgactt[illegible] 839
           |:|||||:||||:|:||||:|||||||:|| ::|||||:|||||||:|||||

Db     821 gtcgacttggaggttgtgcccttgaggcgtggcttccggagctaacgcgttaagtcgacc 880

Qy     840 gcctggggagtacggccgcaaggttaaaactcaaatgaattgacgggggcccgcacaagc 899
           |||:|||||:|||||||||||:|||||:|||:|||||||||||:|||| ||||||:. .

Db     881 gcctggggagtacggccgcaaggttaaaactcaaatgaattgacgggggcccgcacaagc 940

Qy     900 ggtggagcatgtggtttaattcgatgcaacgcgaagaaccttacctactcttgacatccg 959
           ||:|||||:|:||:::||:|||||||||||||||:|||:  :|::||||:|||

Db     941 ggtggagcatgtggtttaattcgaagcaacgcgaagaaccttacctggtcttgacatccg 1000

Qy     960 aaaaacttt[illegible] 1019
             ||| ::: |||||||:|     |:||||:|:|||||  :|||||  |:|:|:,,:|||:

Db    1001 cggaaatttccagagatggattggtgccttcgggaaccggaacacaggtgctgcatggctg 1060

Qy    1020 gtcgtcagctcgtgttgtgaaatgttgggttaagtcccgcaacgagcgcaacccttatcc 1079
           |:||:||||:|:|:|||||||:|:|| :: ||:||:|||||:|  |:|:: |:|:||

Db    1061 gtcgtcagctcgtgtcgtgagatgttgggttaagtcccgcaacgagcgcaacccctatcc 1120

Qy    1080 tttgttgccagcggttaggccgggaactcaaaggagactgccagtgataaactggaggaa 1139
           :::|::|||||||||:  | ||||||||:|||||||||:|||||:|:||||:|||||||

Db    1121 ttgtgttgccagcggttcggccgggaactcaaaggagactgccagtgataaactggaggaa 1180
```

```
Qy    1140 ggtgggatgacgtcaaatcatcatggcccttacgagtagggctacacacgtgctacaat 1199
           ||:|||||: |||:||||:||:|:|||||:|||| |||||:|||||||:||:|| ||
Db    1181 gguggggaugacgucaagucaucauggcccuuacgaccagggcuacacacgugcuacaau 1240

Qy    1200 ggcgcatacaaagagaagcgaccgcgagaggcaagcggacctcataaagtgcgtcgtag 1259
           |||||||:||:|||||||||||||:|||||||||||||||||:||:||||:|||:||||
Db    1241 ggcgcauacaaagagaagcgaccucgcgagaggcaagcggaccucauaaagugcgucguag 1300

Qy    1260 tccggattggagtctgcaactcgactccatgaagtcggaatcgctagtaatcgtagatca 1319
           :||||||::||||:|:|||||:||||:|||:||||:|||||:|||:||:||:|||: ||:||
Db    1301 uccggauuggagucugcaacucgacuccaugaagucggaaucgcuaguaaucguagauca 1360

Qy    1320 gaatgctacggtgaatacgttcccgggccttgtacacaccgcccgtaa 1367
           |||:|| ||||:|||:|||:|||||||||:|:||||||||||||  |
Db    1361 gaaugccacggugaauacguucccgggccuuguacacaccgcccguca 1408


RESULT   5
AAC62270
ID   AAC62270 standard, DNA, 1542 BP.
XX
AC   AAC62270;
XX
DT   19-APR-2001  (first entry)
XX
DE   Escherichia coli reference 16S rDNA sequence
XX
KW   16S rDNA; dechlorinating activity; chlorinated compound; vinyl chloride;
KW   carbontetrachloride; tetrachloroethane; chloroform; dichloromethane;
KW   trichloroethane; dichloroethylene, chlorinating bacteria, ss
XX
OS   Escherichia coli.
XX
PN   WO200063443-A2
XX
PD   26-OCT-2000
XX
PF   13-APR-2000; 2000WO-US09883.
XX
PR   15-APR-1999;   99US-0129511.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PI   Hendrickson ER,  Ebersole RC,
XX
DR   WPI: 2001-024581/03
XX
PT   New 16S rDNA profile derived from Dehalococcoides ethenogenes and
PT   indicative of a dechlorinating bacterial strain -
XX
PS   Disclosure; Page 55; 55pp; English.
XX
CC   The present sequence represents a reference 16S rDNA sequence. The
CC   specification describes 16S rDNA sequences, derived from various strains
CC   of Dehalococcoides ethenogenes. These 16S rDNA profile is linked to
CC   dechlorinating activity  Bacterial strain comprising the 16S rDNA
CC   sequence of the invention are useful for the dechlorination of
CC   chlorinated compounds such as carbontetrachloride, tetrachloroethane,
CC   chloroform, dichloromethane, trichloroethane, dichloroethylene, vinyl
CC   chloride, and chloroaromatics. The 16S rDNA sequence is also useful for
CC   identification of new chlorinating bacteria, as well as sub-typing
CC   strains of Dehalococcoides ethenogenes.
XX
SQ   Sequence 1542 BP; 389 A; 352 C; 487 G; 314 T; 0 other;


  Query Match             86.6%;  Score 1259.6;  DB 22;  Length 1542;
  Best Local Similarity   95.5%;  Pred. No. 0;
  Matches 1306;  Conservative    0;  Mismatches   61;  Indels    1;  Gaps    1;

Qy       1 ggcaggcctaacacatgcaa-tcgagcggcagcggaaagtagcttgctacttt 59
           |||||||||||||||||||| |||| ||| | | | ||| ||:|||| |||||| | ||
```

```
Qy   1140 ggtggggatgacgtcaagtcatcatggcccttacgagtagggctacacacgtgctacaat 1199
          ||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||||||||||
Db   1181 ggtggggatgacgtcaagtcatcatggcccttacggccagggctacacacgtgctacaat 1240

Qy   1200 ggcgcatacaaagagaagcgacctcgcgagagcaagcggacctcataaagtgcgtcgtag 1259
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1241 ggcgcatacaaagagaagcgacctcgcgagagcaagcggacctcataaagtgcgtcgtag 1300

Qy   1260 tccggattggagtctgcaactcgactccatgaagtcggaatcgctagtaatcgtggatca 1319
          |||||| ||||||||||||||||| ||||||||||||||||||||| |
Db   1301 tccggattggagtctgcaactcgactccatgaagtcggaatcgctagtaatcgtggatca 1360

Qy   1320 gaatgctacggtgaatacgttcccgggccttgtacacaccgcccgtca 1367
          |||||| ||||||||||||||||| |||||||||||||||||||| |
Db   1361 gaatgccacggtgaatacgttcccgggccttgtacacaccgcccgtca 1408
```

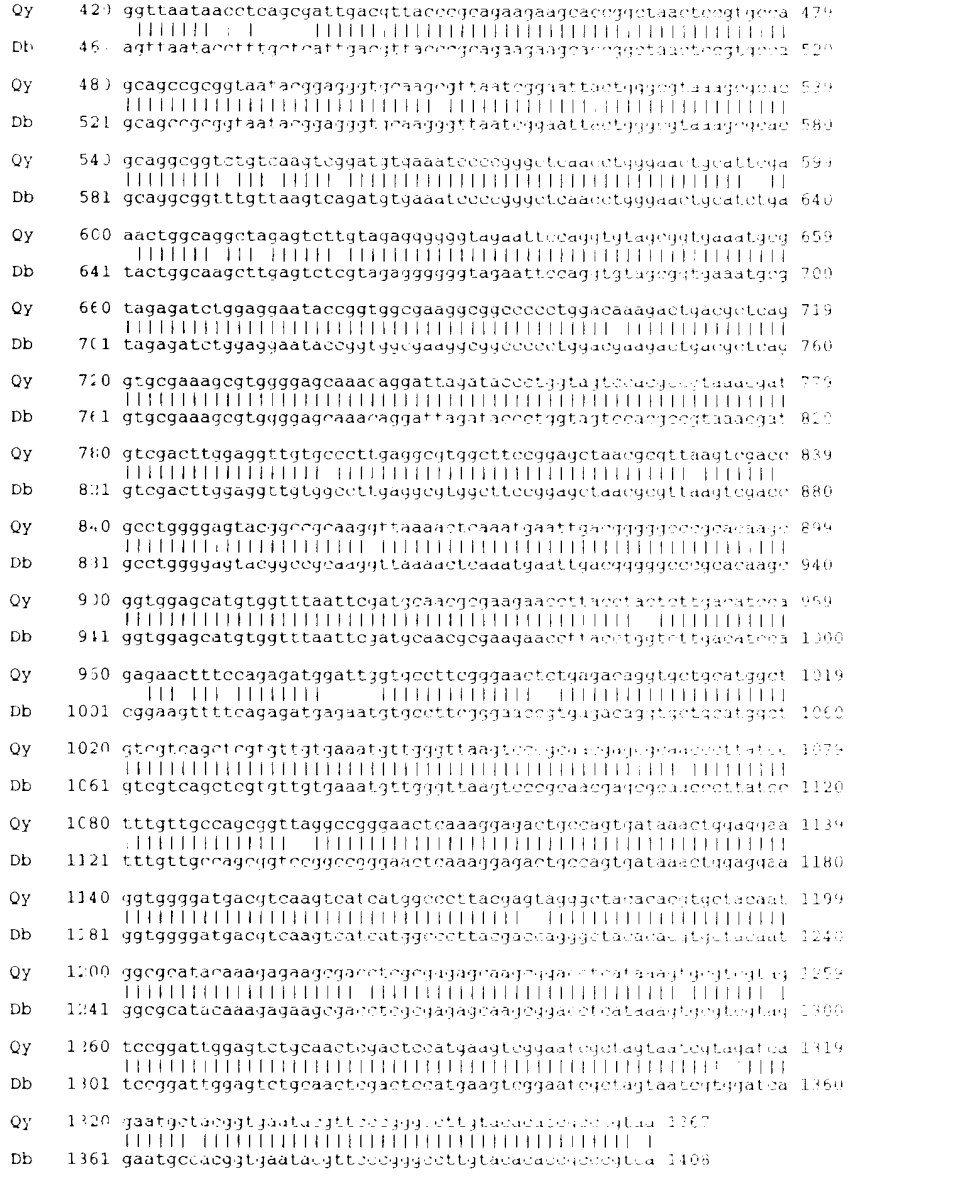
RESULT 5
AAF23015
ID AAF23015 standard; rRNA; 1542 BP.
XX
AC AAF23015;
XX
DT 20-MAR-2001 (first entry)
XX
DE E. coli 16S rRNA sequence.
XX
KW Probe; PCR primer; 5S rRNA; 16S rRNA; 23S rRNA; 28S rRNA; 18S rRNA;
KW Mycobacterium, Enterococcus, Chlamydia, Mycoplasma, E. coli, Legionella,
KW Salmonella, Pseudomonas, Campylobacter, Neisseria gonorrheae; fungus;
KW bacterium; ss.
XX
OS Escherichia coli.
XX
PN US6150517-A.
XX
PD 21-NOV-2000.
XX
PF 30-MAY-1995; 95US-0454063.
XX
PR 22-FEB-1994; 94US-0200866.
PR 24-NOV-1987; 87US-0295208.
PR 24-NOV-1987; 87WO-US03009.
PR 11-DEC-1991; 91US-0806929.
PR 24-NOV-1986; 86US-0934244.
PR 07-AUG-1987; 87US-0083542.
XX
PA (GENP-) GEN-PROBE INC.
XX
PI McDonough SH, Kop JA, Smith RD, Hogan JJ;
XX
DR WPI; 2001-060029/07.
XX
PT Preparing a probe for nucleic acid hybridization assays comprises
PT constructing a nucleotide polymer sufficiently complementary to
PT hybridize to an rRNA region that distinguishes non-viral target from
PT non-viral non-target species -
XX
PS Disclosure; Fig 1, 75pp, English.
XX
CC The present invention provides novel methods of producing probes for use
CC in the identification of a number of microorganisms. These include E.
CC coli, Mycobacteria, Mycoplasma, Campylobacter, Chlamydia, Enterobacter,
CC Legionella, Salmonella, Pseudomonas, Neisseria gonorrheae, fungi and
CC bacteria.
XX
SQ Sequence 1542 BP, 389 A, 351 C, 488 G, 314 U, 0 Other,

Query Match 86.5%; Score 1258; DB 22; Length 1542;
Best Local Similarity 76.7%; Pred. No. 0;
Matches 1049; Conservative 256; Mismatches 62; Indels 1; Gaps 1;

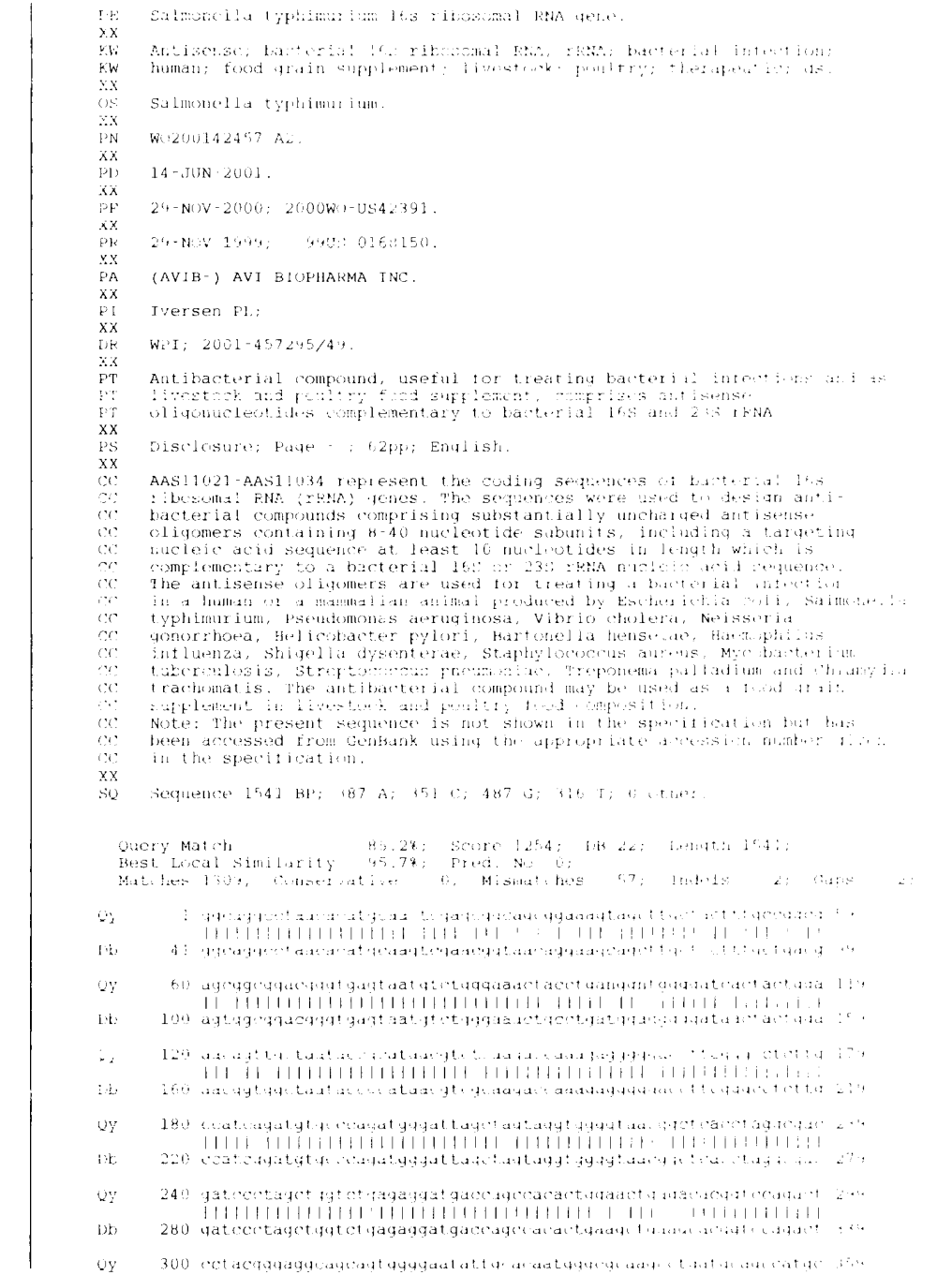
```
Qy      1 ggcgaggctaacacatgcaagtcgaacggcagcacgaaaagtagcttgctactttgcgg 60
          ||||||||:||||||:|||| : || ||| | | | ||| :|:::|: |:::|: : ||
Db     41 ggcgagcgccgaacacgatgaacggagcggaaacagtaagaagcttgcttctcgctgcg 100

Qy     61 agtggcggacgggtgagtaatgtctgggaaactgcctgatggagggggataactactgga 120
          || |||||||||:|||:|||:|:|:|||||||: ||:|| ||  |||||: |||:|:||
Db    101 agcgggcgatagtgtgatgaatatctggaaagccgccttgagtggggataacgtacagga 160

Qy    121 aacagttggtaataccgcatacgtcctacgggagccaaagagggggacccttcgggcctct 180
          ||| |: ||:||:|||||:||||:| ||||||||||||||||||||||:|||||||:|:|
Db    161 aacgggtaccgcaatgcgtgatgaggctgaaagcgggggaaccttcgcgacctcgcgctg 220

Qy    181 cgcatcagatatgcctagttggtgaggtaatggctcaccaaggcgacgatccgtaactgac 240
          |||:| ||:|:|||||:|||:::|||:|||:|:|:|||||||||||||||:|||||||||
Db    221 ccaacggaaggagccccagaagtggaaagcgaaggcagcttgctgcgcggtatccagaagc 280

Qy    241 gatcctagctggtctgagaggatgaccagccacactggaactgagacacggtccagactcc 300
          ||:|||:|||:||:|||||||:|||||||||||||:|||||:|||||||| ||:  |:::
Db    281 gatccgacgcggagcgagaggaagagcgcgacgcgcgcgcgcgacgagcgcgcgcgcgcg 340

Qy    301 cctacgggaggcagcagtggggaatattgcacaatgggcgcaagcctgatgcagccatgcc 360
          ||:|||||||||||||:|||||:|:||||||:|||||||||||||: |:|||||||:||
Db    341 ccgacgggagcgcagcagtgcggaaatagtgcacaatgggcgcaagcctgacgcagcgatgc 400

Qy    361 cgcgtgtatgaagaaggccttcgggttgtaaagtactttcagcggggaggaagggcgttga 420
          ||||:|:|:|||||||||||:|||:|:|||:|||:|||||||||||. |||||  : :
Db    401 cgcgtgcgcgaagaaggccgtccgtgcgttgtaaagcgcttttgccagcgaggacgggcgacg 460

Qy    421 ggttaatagcttgtcgtgattgacgttactcgcagaagaagcaccggctaactccgtgcca 480
          |::|:| | :    |:|||:|:|||||||||||||||||||||||||:|||:|:|:.|||
Db    461 agcgaagcgccgttgccgcgcgagcgaagcgacgcagcgcgcgcgcgcgcgcgcgcgcca 520

Qy    481 gcagccgcggtaatacggagggtgcaagcgttaatcggaattactgggcgtaaagcgcgcg 540
          |||||||||||:|:|||||:|||||||:|:|||||:|:|||||:|:|:|:|:|:|:|||
Db    521 gcgagccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 580

Qy    541 gcaggcggtttgttaagtcagatgtgaaatccccgggctcaacctgggaactgcatctgat 600
          |||||||: :|: |||:| |:|:|||||||||||||||:|||||||||||||||||||||
Db    581 gcaggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 640

Qy    601 actgggcgagctagagtatggtagagggtggtggaatttcctgtgtagcggtgaaatgcgt 660
          ||:|||| ||: |||:|: |:|||||||:||||:|||||:|||||:|||||:|||||||||
Db    641 gcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 700

Qy    661 agagatctaggaggaataccggtggcgaaggcggccccctggacgaagactgacgctcagg 720
          :||||||:|||||||:|||||:|||||||||||||||||||||||||||   |||||||:||
Db    701 gcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 760

Qy    721 gtgcgaaagcgtggggagcaaacaggattagataccctggtagtccacgccgtaaacgatg 780
          |:|||||||||:|||||  |||||||:|||:||||:||||:|||||||||||||||||||
Db    761 gcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 820

Qy    781 tcgaacttggaggttgtgcccttgaggcgtggcttccggagctaacgcgttaagttcgccg 840
          |:|||||:|||||:|:||||:|||||||:|||||||||||||||:|||:|||||||||||
Db    821 gcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 880

Qy    841 gcctggggagtacggccgcaaggttaaaactcaaatgaattgacgggggcccgcacaagcgg 900
          |||:|||||:|:|||||||||:|||||:|:|:|||:|||||||||||||||||||||||||
Db    881 gcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 940

Qy    901 ggtggagcatgtggtttaattcgatgcaacgcgaagaaccttacctactcttgacatccag 960
          ||:|||||:|:|:::|::|||:|||||||||||||||||:|||:  ::::|:|:|||
Db    941 ggtggagcatgcggttaattcgacgcaacgcgcagaaccttaccaacccttgacatccag 1000

Qy    961 agaagcttttccagagatggattggtgccttcgggaactgtctgacacaggtgctgcatgg 1020
            ||| ::: |||||||:     |:|||::|||||||  :|||||||||:|:|||:|:
Db   1001 cggaagttttcagagatgagaattgtgccttagtggaacgcgagacaggtgctgcatgg 1060

Qy   1021 ctcgtcagctcgtgttgtgaaatgttgggttaagtcccgcaacgagcgcaaccctttatcc 1080
          |:|:|||||:|:|:|||||||||||||||||||||||||||||||||||  :|||||:|:||
Db   1061 gtcgtcagctcgtgtcgtgagatgttcggttaagtccggcaacgagcgcaacccctgtccc 1120
```

```
Qy   1080 ttgttgccagcggttaggccgggaactcaaaggagactgccagtgataaactggaggaa 1139
          ::::::|||||||||:  ||||||||||:||||||||||:|||| :|||:||||:|||||||
Db   1121 uuuguugccagcggucccggccgggaacucaaaggagacugccagugauaaacuggaggaa 1180

Qy   1140 ggtggggatgacgtcaagtcatcatggcccttacgagtagggctacacacgtgctacaat 1199
          ||:|||||:||||:||||:||:||:|||||::||||  |||||:|||||||:|:|||||:
Db   1181 gguggggaugacgucaagucaucauggcccuuacgaccagggcuacacacgugcuacaau 1240

Qy   1200 ggcgcatacaaagagaagcgacctcgcgagagcaagcggacctcataaagtgcgtcgtag 1259
          ||||||:|||||||:|||||||:|||||||||||||||||:||:||||:|||||||||||
Db   1241 ggcgcauacaaagagaagcgaccucgcgagagcaagcggaccucauaaagugcgucguag 1300

Qy   1260 tccggattggagtctgcaactcgactccatgaagtcggaatcgctagtaatcgtagatca 1319
          :||||||:||||:|:||||:||||:|||:||||:||||:|||:||:||:||:|:: ||:||
Db   1301 uccggauuggagucugcaacucgacuccaugaagucggaaucgcuaguaaucguggauca 1360

Qy   1320 gaatgctacggtgaatacgttcccgggccttgtacacaccgcccgtaa 1367
          |||:|| |||||:|||:|||::|| ||||||:|:|||||||||||| |
Db   136. gaaugccacggugaauacguucccgggccuuguacacaccgcccguca 1408


RESULT   7
AAX24983
ID   AAX24983 standard; DNA; 5097 BP.
XX
AC   AAX24983;
XX
DT   05-JUL-1999  (first entry)
XX
DE   E. coli MG1655 rrnA operon (16S-spacer-23S-spacer-5S).
XX
KW   Speciation; ribotyping; species discrimination; marker; 16S;
KW   restriction fragment length polymorphism; bacterium; fungus;
KW   pathogen; rrnA operon; 16S RNA gene; 23S RNA gene; ds.
XX
OS   Escherichia coli.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1..1547
FT                   /*tag= a
FT                   /label= 16S
FT   misc_feature    1980..4884
FT                   /*tag= b
FT                   /label= 23S
FT   misc_feature    4978..5097
FT                   /*tag= c
FT                   /label= 5S
XX
PN   WO9905325-A1.
XX
PD   04-FEB-1999.
XX
PF   24-JUL-1998;   98WO-US15464.
XX
PR   25-JUL-1997;   97US-0053097.
XX
PA   (UYBO-) UNIV BOSTON.
XX
PI   Goldstein RN;
XX
DR   WPI; 1999-142969/12.
XX
PT   Determining species of bacteria and fungi - useful for
PT   distinguishing between bacterial/fungal species, and for determining
PT   the identity of bacterial/fungal pathogens in biological samples
XX
PS   Disclosure; Fig 7 (18/67 21/67); 133pp; English.
XX
CC   This is the DNA sequence of the Escherichia coli strain MG1655
CC   rrnA operon (16S-spacer-23S-spacer-5S).  Restriction sites for
CC   enzymes cutting the operon 5 times or less have been determined.
CC   E. coli rrnA-rrnH operon sequences are provided (see AAX24985-89).
CC   Methods and compositions are described for determining the species
CC   of an unknown bacterium or fungus in a sample.  The method involves
CC   isolating and digesting bacterial (or fungal) DNA encoding 16S and
CC   23S rRNA from a sample with restriction enzymes, detecting the
CC   products, and comparing them to signature bands from a number of
CC   bacteria.  The method generates a species conserved set of RFLP
CC   bands, unique for each species.  These species-conserved sets
CC   represent precise markers appropriate for inter-species
CC   discriminatory purposes (i.e. to determine the species of a given,
CC   unknown isolate e.g. in a clinical specimen).  In contrast to
CC   conventional ribotyping, the present invention utilises the
CC   ribosomal operon sequences which vary less than 3% (and more
CC   preferably less than 2%) within a species and vary between species.
CC   The method is useful for medical, food, agricultural and
CC   environmental testing.  It does not require sequencing of nucleic
CC   acid from biological samples.
XX
SQ   Sequence 5097 BP; 1333 A; 1151 C; 1565 G; 1068 T; 0 other;

  Query Match             86.4%;  Score 1256.4;  DB 20;  Length 5097;
  Best Local Similarity   95.3%;  Prod. No. 0;
  Matches 1304;  Conservative    0;  Mismatches   63;  Indels    1;  Gaps    1;

Qy       1 ggcaggcctaacacatgcaa-tcgagcggcagcgggaaagtagcttgctactttgccggcg 59
           |||||||||||||||||||| |||| |||   | | ||| ||||||| :|||| | ||
Db      41 ggcaggcctaacacatgcaagtcgaacggtaacaggaagaagcttgcttctttgctgacgt 100

Qy      60 agcggcggacgggtgagtaatgtctgggaaactgcctgatggagggggataactactggaa 119
           || |||||||||||||||||||||||||||| | |||| ||  | ||| ||||||||||
Db     101 agtggcgaacgggtgagtaatatgtatgggaaactgcctgatggagggggataactactgga 160

Qy     120 aacggtagctaataccgcataacgtcgcaagaccaaagagggggaccttcgggcctctttgg 179
           ||| || ||||||||||| |||||| ||||||||||||||||||| |||||||||||||||
Db     161 aacattatgctaataccatataagcgcacaaccgaaagatttggtcttttcggagcttttgg 220

Qy     180 agattagctatgtcgccagatgagcctaggtcggattagctagttggtgaggtaatggctca 239
           ||||| |||||||||||||||||||||||||||||||||||||||||||||| |||||||||
Db     221 agattaggatgttggccagatgagcctaggtcggattagctagttggtgaggtaatggctca 280

Qy     240 ccaaggcgacgatccgtaactggtctgagaggatgatcagtcacactggaactgagacacggt 299
           |||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     281 ccaaggcgacgatccgtaactggtctgagaggatgatcagtcacactggaactgagacacggt 340

Qy     300 ccagactcctacgggaggcagcagtggggaatattgcacaatgggcgcaagcctgatgcagc 359
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     341 ccagactcctacgggaggcagcagtggggaatattgcacaatgggcgcaagcctgatgcagc 400

Qy     360 cgccgtgtatgaagaaggccttcgggttgtaaagtactttcagcggggaggaagggagtaaa 419
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     401 cgccgtgtatgaagaaggccttcgggttgtaaagtactttcagcggggaggaagggagtaaa 460

Qy     420 gttaatacctttgctcattgacgttacccgcagaagaagcaccggctaactccgtgccagc 479
           ||||||| | |     ||||||||||||||||||||||||||||||||||||||||||||||
Db     461 agttaatacctttgctcattgacgttacccgcagaagaagcaccggctaactccgtgccagc 520

Qy     480 agccgcggtaatacgcagggtgcgagcgttaatcggaattactgggcgtaaagcgcgcgc 539
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     521 agccgcggtaatacgcagggtgcaagcgttaatcggaattactgggcgtaaagcgcgcgc 580

Qy     540 gcaggcggttctttaagtctgatgtgaaagcccgggctcaacctgggaactgcattggaaa 599
           |||||||| ||| ||| | |||||||||||||||||||||||||||||||||| |||| ||
Db     581 gcaggcggtttgttaagtcagatgtgaaatccccgggctcaacctgggaactgcatctgata 640

Qy     600 actggcaagctagagtcttgtagaggggggtagaattccaggtgtagcggtgaaatgcgta 659
           ||||||| ||| |||||| ||||||||||||||||||||||||||||||| ||| |||||
Db     641 tactggcaagcttgagtctcgtagaggggggtagaattccaggtgtagcggtgaaatgcgta 700

Qy     660 tagaggatctggaggaatacc-ggtggcgaaggcggccccctggacgaagactgacgctcag 718
           |||||||||||||||||||| |||||||||||||||||||| |||||| |||||| |||||
Db     701 tagaggatctggaggaataccagtggcgaaggcggccccctgaacgaagactgacgctcag 760

Qy     719 gtgcgaaagcgtggggagcaaacaggattagataccctggtagtccacgccgtaaacgatgt 778
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     761 gtgcgaaagcgtggggagcaaacaggattagataccctggtagtccacgccgtaaacgatgt 820
```

```
Db     761 gtgcgaaagcgtggggagcaaacaggattagataccctggtagtccacgccgtaaacgat 820

Qy     780 gtcgacttggaggttgtgcccttgaggcgtggcttccggagctaacgcgttaagtcgacc 839
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     821 gtcgacttggaggttgtgcccttgaggcgtggcttccggagctaacgcgttaagtcgacc 880

Qy     840 gcctggggagtacggccgcaaggttaaaactcaaatgaattgacgggggcccgcacaagc 899
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     881 gcctggggagtacggccgcaaggttaaaactcaaatgaattgacgggggcccgcacaagc 940

Qy     900 ggtggagcatgtggtttaattcgatgcaacgcgaagaaccttacctactcttgacatcca 959
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     941 ggtggagcatgtggtttaattcgatgcaacgcgaagaaccttacctggtcttgacatcca 1000

Qy     960 gagaactttccagagatggattgatgccttcgggaactctgaggcaggtgctgcatggct 1019
           ||| ||| |||||||     ||||||||||||  ||||||||||||||||||||||||
Db    1001 cggaagttttcagagatcagaatgtgccttcgggaaccgtgaggcaggtgctgcatggct 1060

Qy    1020 gtcgtcagctcgtgttgtgaaatgttgggttaagtcccgcaacgagcgcaacccttatcc 1079
           |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db    1061 gtcgtcagctcgtgttgtgaaatgttgggttaagtcccgcaacgagcgcaacccttatcc 1120

Qy    1080 tttgttgccagcggttaggccgggaactcaaaggagactgccagtgataaactggaggaa 1139
           ||||||||||||  |||||| |||||||||||||||||||||||||||||| ||||||
Db    1121 tttgttgccagcggtccggccgggaactcaaaggagactgccagtgataaactggaggaa 1180

Qy    1140 ggtggggatgacgtcaagtcatcatggcccttacggagtagggctacacacgtgctacaat 1199
           ||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db    1181 ggtggggatgacgtcaagtcatcatggcccttacgaccagggctacacacgtgctacaat 1240

Qy    1200 ggcgcatacaaagagaagcgacctcgcgagagcaagcggacctcataaagtgcgtcgtag 1259
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1241 ggcgcatacaaagagaagcgacctcgcgagagcaagcggacctcataaagtgcgtcgtag 1300

Qy    1260 tccggattggagtctgcaactcgactccatgaagtcggaatcgctagtaatcgtagatca 1319
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1301 tccggattggagtctgcaactcgactccatgaagtcggaatcgctagtaatcgtggatca 1360

Qy    1320 gaatgctacggtgaatacgttcccgggccttgtacacaccgcccgtcaa 1367
           |||||| ||||||||||||||||||||||||||||||||||||||||| |
Db    1361 gaatgccacggtgaatacgttcccgggccttgtacacaccgcccgtca 1408


RESULT   8
AAX24984
ID   AAX24984 standard; DNA; 5098 BP.
XX
AC   AAX24984;
XX
DT   05-JUL-1999  (first entry)
XX
DE   E. coli MG1655 rrnB operon (16S-spacer 23S spacer 5S).
XX
KW   Speciation; ribotyping; species discrimination; marker; RFLP;
KW   restriction fragment length polymorphism; bacterium; fungus;
KW   pathogen, rrnB operon, 16S RNA gene, 23S RNA gene, 5S.
XX
OS   Escherichia coli.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1..1542
FT                   /*tag=  a
FT                   /label= 16S
FT   misc_feature    1983..4884
FT                   /*tag=  b
FT                   /label= 23S
FT   misc_feature    4979..5098
FT                   /*tag=  c
FT                   /label= 5S
XX
PN   WO9905325-A1.
XX
PD   04-FEB-1999.
XX
PF   24-JUL-1998;    98WO-CS15464.
XX
PR   25-JUL-1997;    97US-0053097.
XX
PA   (UYBO-) UNIV BOSTON.
XX
PI   Goldstein RN;
XX
DR   WPI; 1999-142969/12.
XX
PT   Determining species of bacteria and fungi - useful for
PT   distinguishing between bacterial/fungal species, and for determining
PT   the identity of bacterial/fungal pathogens in biological samples
XX
PS   Disclosure; Fig 7 (25/67-28/67); 133pp; English.
XX
CC   This is the DNA sequence of the Escherichia coli strain MG1655
CC   rrnB operon (16S spacer 23S spacer 5S). Restriction sites for
CC   enzymes cutting the operon 5 times or less have been determined.
CC   E. coli rrnA rrnB operon sequences are provided (see AAX24985-89).
CC   Methods and compositions are described for determining the species
CC   of an unknown bacterium or fungus in a sample. The method involves
CC   isolating and digesting bacterial (or fungal) DNA encoding 16S and
CC   23S rRNA from a sample with restriction enzymes, detecting the
CC   products, and comparing them to signature bands from a number of
CC   bacteria. The method generates a species-conserved set of RFLP
CC   bands, unique for each species. These species-conserved sets
CC   represent precise markers appropriate for inter-species
CC   discriminatory purposes (i.e. to determine the species of a given
CC   unknown isolate e.g. in a clinical specimen). In contrast to
CC   conventional ribotyping, the present invention utilises the
CC   ribosomal operon sequences which vary less than 3% (and more
CC   preferably less than 2%) within a species and vary between species.
CC   The method is useful for medical, food, agricultural and
CC   environmental testing. It does not require sequencing of nucleic
CC   acid from biological samples.
XX
SQ   Sequence 5098 BP; 1344 A; 1151 C; 1552 G; 1061 T; 0 other;

  Query Match             86.4%;  Score 1256.4;  DB 20;  Length 5098;
  Best Local Similarity   95.3%;  Prod. No. 0;
  Matches 1304;  Conservative    0;  Mismatches   64;  Indels    1;  Gaps    1

Qy       1 ggcaggcctaacacatgcaag-tcgaacggcagcacgggaagtagcttgctactttgccggcg 59
           |||||||||||||||||||| |||| ||| | | | ||| ||||| | ||||| | |||| |
Db      41 ggcaggcctaacacatgcaagtcgaacggtaacaggaagaagcttgcttctttgctgacgag 100

Qy      60 agcggcggacgggtgagtaatgtctgggaaactgcctgatggagggggataactactggaaa 119
           || |||||||||||||| ||||||.|||||| |||| ||  ||  |  ||  |  |  |  |
Db     101 agtggcggacgggtgagtaatgtctgggaaatgccctgatggagggggataactactggtta 160

Qy     120 gattagttgctaataccgcatacgtcctacgggagaaagcaggggaccttcgggccttgcgc 179
           ||| || ||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db     161 gattagatagtcgatacc.ctataggt ccttt.aaggaaagatttatcgattggggatgagc 219

Qy     180 ccatcagatgtgcccagatgggattagctagtaggtggggtaacggctcacctaggcgacgat 239
           ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     220 ccatcggatgtgcccagatgggattagctagtaggtggggtaacggctcacctaggcgacgat 279

Qy     240 gatccctagctggtctgagaggatgaccagccacactggaactgagacacggtccagactcct 299
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     280 gatccctagctggtctgagaggatgaccagccacactggaactgagacacggtccagactcct 339

Qy     300 acgggaggcagcagtggggaatattgcacaatgggcgcaagcctgatgcagccatgccgcgt 359
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     340 acgggaggcagcagtggggaatattgcacaatgggcgcaagcctgatgcagccatgccgcgt 399

Qy     360 cagtgtatgaaggaagccttcgggttgtaaagtactttcagcggggaggaagggagtaaagt 419
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     400 cagtgtatgaaggaagccttcgggttgtaaagtactttcagcggggaggaagggagtaaagt 459
```



RESULT 9
AAS11022
ID AAS11022 standard; DNA; 1541 BP.
XX
AC AAS11022;
XX
DT 24-OCT-2001 (first entry)
XX



DE Salmonella typhimurium 16s ribosomal RNA gene.
XX
KW Antisense; bacterial 16S ribosomal RNA; rRNA; bacterial infection;
KW human; food grain supplement; livestock; poultry; therapeutic; ds.
XX
OS Salmonella typhimurium.
XX
PN WO200142457 A2.
XX
PD 14-JUN-2001.
XX
PF 29-NOV-2000; 2000WO-US42391.
XX
PR 29-NOV-1999; 99US-0168150.
XX
PA (AVIB-) AVI BIOPHARMA INC.
XX
PI Iversen PL;
XX
DR WPI; 2001-457295/49.
XX
PT Antibacterial compound, useful for treating bacterial infections and as
PT livestock and poultry feed supplement, comprises antisense
PT oligonucleotides complementary to bacterial 16S and 23S rRNA
XX
PS Disclosure; Page - ; 62pp; English.
XX
CC AAS11021-AAS11034 represent the coding sequences of bacterial 16s
CC ribosomal RNA (rRNA) genes. The sequences were used to design anti-
CC bacterial compounds comprising substantially uncharged antisense
CC oligomers containing 8-40 nucleotide subunits, including a targeting
CC nucleic acid sequence at least 10 nucleotides in length which is
CC complementary to a bacterial 16S or 23S rRNA nucleic acid sequence.
CC The antisense oligomers are used for treating a bacterial infection
CC in a human or a mammalian animal produced by Escherichia coli, Salmonella
CC typhimurium, Pseudomonas aeruginosa, Vibrio cholera, Neisseria
CC gonorrhoea, Helicobacter pylori, Bartonella henselae, Haemophilus
CC influenza, Shigella dysenteriae, Staphylococcus aureus, Mycobacterium
CC tuberculosis, Streptococcus pneumoniae, Treponema pallidum and Chlamydia
CC trachomatis. The antibacterial compound may be used as a food grain
CC supplement in livestock and poultry feed compositions.
CC Note: The present sequence is not shown in the specification but has
CC been accessed from GenBank using the appropriate accession number given
CC in the specification.
XX
SQ Sequence 1541 BP; 487 A; 451 C; 487 G; 416 T; 0 other;

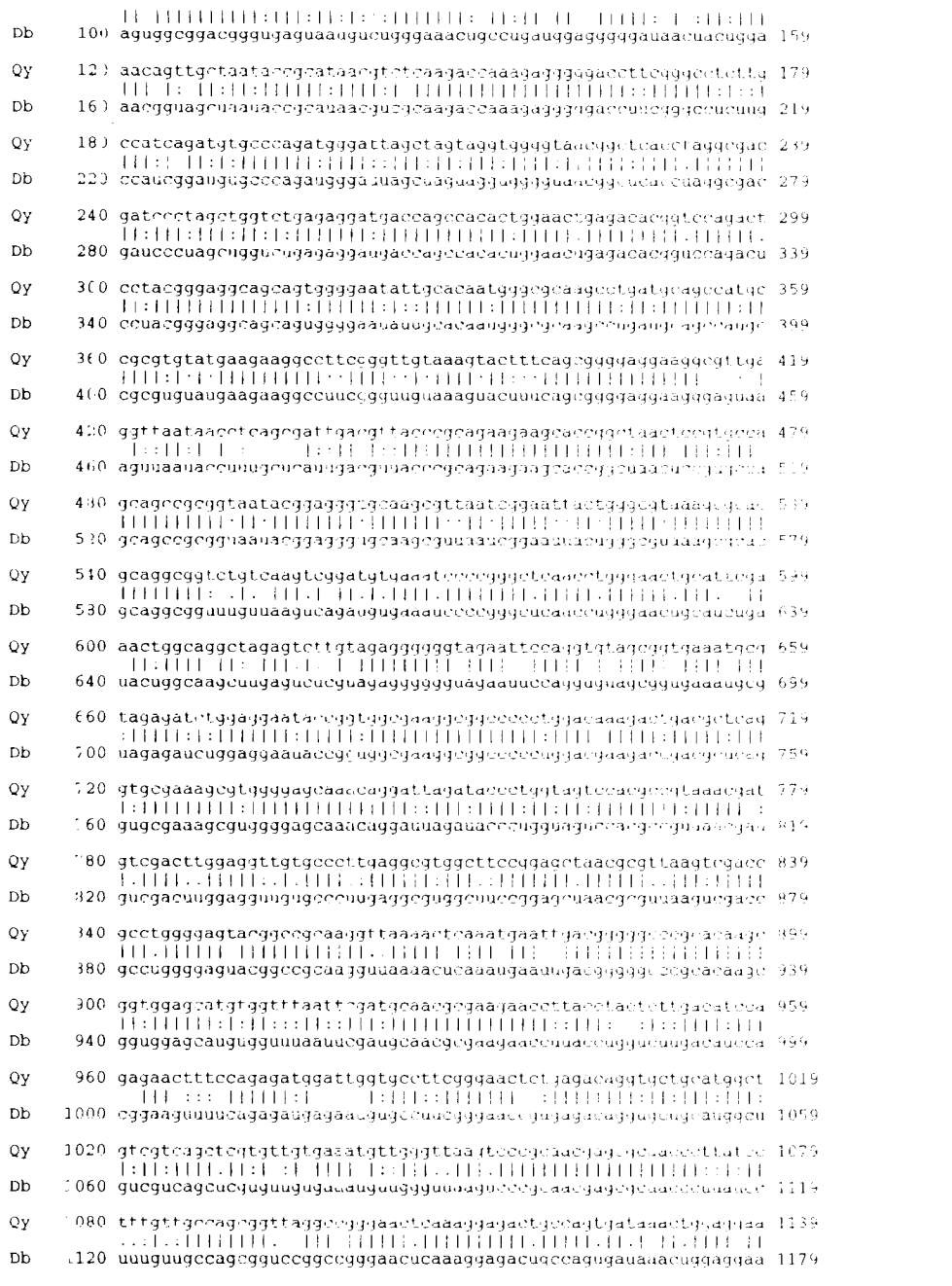
Query Match 85.2%; Score 1254; DB 22; Length 1541;
Best Local Similarity 95.7%; Pred. No. 0;
Matches 1304; Conservative 0; Mismatches 57; Indels 2; Gaps 2;

```
Db    340 cctacgggaggcagcagtgggaatattgcacaatgggcgcaagcctgatgcagccatgc 399

Qy    360 cgcgtgtatgaagaaggccttcgggttgtaaagtactttcagcggggaggaagggattga 419
          ||||||||||| |||||||||||||| |||||||||||||||||||||||||||||   ||||
Db    400 cgcgtgtatgaagaaggccttcggqttgtaaagtactttcagcgggqaggaagggtatttgt 459

Qy    420 ggttaataacctcagcgattgacgttacccgcagaagaagcaccggctaactccgtgccag 479
          ||||||||||| |||| ||||||||||||||||||||||||||||||||||||||||||||
Db    460 ggttaataacccgcagcaattgacgttacccgcagaagaagcaccggctaactccgtgccag 519

Qy    480 gcagccgcggtaatacggagggtgcaagcgttaatcggaattactgggcgtaaagcgcgcg 539
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    520 gcagccgcggtaatacggagggtgcaagcgttaatcggaattactgggcgtaaagcgcgcg 579

Qy    540 gcaggcggtctgttaagtcggatgtgaaatccccgggctcaacctgggaactgcattcga 599
          ||||||||| ||| ||||| ||||||||||||||||||||||||||||||||||| ||
Db    580 gcaggcggtttgttaagtcagatgtgaaatccccgggctcaacctgggaactgcatctga 639

Qy    600 aactggcaggctagagtcttgtagaggggggtagaattccaggtgtagcggtgaaatgcg 659
          ||||||| ||| |||||| ||||||||||||||||||||||||||||||||||||||||
Db    640 tactggcaagcttgagtctcgtagaggggggtagaattccaggtgtagcggtgaaatgcg 699

Qy    660 tagagatctggaggaataccggtggcgaaggcggccccctggacgaagactgacgctcag 719
          |||||||||||||||||| |||||||||||||||||||| ||||||| |||||||
Db    700 tagagatctggaggaataccggtggcgaaggcggccccctggacgaagactgacgctcag 759

Qy    720 gtgcgaaagcgtggggagcaaacaggattagataccctggtagtccacgccgtaaacgat 779
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    760 gtgcgaaagcgtggggagcaaacaggattagataccctggtagtccacgccgtaaacgat 819

Qy    780 gtcgacttggaggttgtgcccttgaggcgtggcttccggagctaacgcgttaagtcgacc 839
          || ||||||||||||||||||||||||||||||||||||||| ||||||||||| ||
Db    820 gtctacttggaggttgtgcccttgaggcgtggcttccggagctaacgcgttaagtagacc 879

Qy    840 gcctggggagtacggccgcaaggttaaaactcaaatgaattgacgggggcccgcacaagc 899
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    880 gcttggggagtacggccgcaagcttaaaactcaaatgaattgacgggggcccgcacaagc 939

Qy    900 ggtggagcatgtggtttaattcgatgcaacgcgaagaaccttacctactcttgacatcca 959
          ||||||||||||||||||||||||||||||||||||||||||||  ||||||||||||
Db    940 ggtggagcatgtggtttaattcgatgcaacgcgaagaaccttacctggtcttgacatcca 999

Qy    960 gagaactttccagagatggattggtgccttcgggaactctgagacaggtgctgcatggct 1019
          ||||||||||||||||   | ||||||||||||||| ||||||||||||||||||||||
Db   1000 cagaacttccagagatgagattcgtgccttcgggaactgtgagacaggtgctgcatggct 1059

Qy   1020 gtcgtcagctcgtgttgtgaaatgttgggttaagtcccgcaacgagcgcaaccctttatcc 1079
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1060 gtcgtcagctcgtgttgtgaaatgttgggttaagtcccgcaacgagcgcaaccctttatcc 1119

Qy   1080 tttgttgccagcggttaggccgggaactcaaaggagactgccagtgataaactggaggaaa 1139
          ||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||||
Db   1120 tttgttgccagcggtccggccgggaactcaaaggagactgccagtgataaactggaggaaa 1179

Qy   1140 ggtggggatgacgtcaagtcatcatggcccttacgagtagggctacacacgtgctacaat 1199
          ||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db   1180 ggtggggatgacgtcaagtcatcatggcccttacgaccagggctacacacgtgctacaat 1239

Qy   1200 ggcgcatacaaagagaagcgacctcgcgagagcaagcggacctcataaagtgcgtcgtag 1259
          |||||||||||||||||||| |||||||||||||||||||||||| ||||||| |
Db   1240 ggcgcatacaaagagaagcgacctcgcgagagcaagcggacctcataaagtgcgtcgtag 1299

Qy   1260 tccggattggagtctgcaactcgactccatgaagtcggaatcgctagtaatcgtggatca 1319
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1300 tccggattggagtctgcaactcgactccatgaagtcggaatcgctagtaatcgtggatca 1359

Qy   1320 gaatgctacggtgaatacgttcccgggccttgtacacaccgcccgtaa 1367
          |||||| |||||||||||||||||||||||||||||||||||||| |
Db   1360 gaatgccacggtgaatacgttcccgggccttgtacacaccgcccgtca 1407
```

. RESULT 10

```
AAX24987
ID   AAX24987 standard; DNA; 5014 BP.
XX
AC   AAX24987;
XX
DT   05-JUL-1999  (first entry)
XX
DE   E. coli M55055 rrnE operon (16S spacer 23S spacer 5S).
XX
KW   Speciation; ribotyping; species discrimination; marker; RFLP;
KW   restriction fragment length polymorphism; bacterium; fungus;
KW   pathogen; rrnE operon; 16S RNA gene; 23S RNA gene; 5S.
XX
OS   Escherichia coli.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1..1547
FT                   /*tag= a
FT                   /label= 16S
FT   misc_feature    1797..4819
FT                   /*tag= b
FT                   /label= 23S
FT   misc_feature    4895..5014
FT                   /*tag= c
FT                   /label= 5S
XX
PN   WO9905325-A1.
XX
PD   04-FEB-1999.
XX
PF   24-JUL-1998;   98WO-US15464.
XX
PR   25-JUL-1997;   97US-0053097.
XX
PA   (UYBO-) UNIV BOSTON.
XX
PI   Goldstein RN;
XX
DR   WPI; 1999-142969/12.
XX
PT   Determining species of bacteria and fungi - useful for
PT   distinguishing between bacterial/fungal species, and for determining
PT   the identity of bacterial/fungal pathogens in biological samples
XX
PS   Disclosure; Fig 7 (46/67 49/67); 133pp; English.
XX
CC   This is the DNA sequence of the Escherichia coli strain M55055
CC   rrnE operon (16S spacer-23S spacer 5S). Restriction sites for
CC   enzymes cutting the operon 5 times or less have been determined.
CC   E. coli rrnA-rrnH operon sequences are provided (see AAX24983-AAX).
CC   Methods and compositions are described for determining the species
CC   of an unknown bacterium or fungus in a sample. The method involves
CC   isolating and digesting bacterial (or fungal) DNA encoding 16S and
CC   23S rRNA from a sample with restriction enzymes, detecting the
CC   products, and comparing them to signature bands from a number of
CC   bacteria. The method generates a species conserved set of RFLP
CC   bands, unique for each species. These species-conserved sets
CC   represent precise markers appopropriate for inter-species
CC   discriminatory purposes (i.e. to determine the species of a given
CC   unknown isolate e.g. in a clinical specimen). In contrast to
CC   conventional ribotyping, the present invention utilises the
CC   ribosomal operon sequences which vary less than 1% (and more
CC   preferably less than 2%) within a species and vary between species.
CC   The method is useful for medical, food, agricultural and
CC   environmental testing. It does not require sequencing of nucleic
CC   acid from biological samples.
XX
SQ   Sequence 5014 BP; 1408 A; 1129 C; 1537 G; 1040 T; 0 other;

  Query Match             86.1%;  Score 1253.2;  DB 2;  Length 5014;
  Best Local Similarity   95.2%;  Pred. No. 0;
  Matches 1302,  Conservative    0,  Mismatches   65;  Indels    1;  Gaps    1;
```

```
Qy      60 agcggcggacgggtgagtaatgtc gggaaactacctgatggatgggggataactactgaga 119
           || ||||||||||||||||||||||||||||||||| ||||| ||  |||||| ||||||||
Db     101 agtggcggacgggtgagtaatgtctgggaaactgcctgatggagggggataactactggaa 160

Qy     120 aacagttgctaataccgcataacgtctcaagaccaaagagggggaccttcgggcctcttg 179
           ||| || |||||||| |||||||| ||||||||||||||||||||||||||||||||||
Db     161 aacggtagctaatacctcataacgtcgcaagaccaaagagggggaccttcggcctcttg 220

Qy     180 ccatcagatgtgcccagatgggattagctagtaggtggggtaacggctcacctaggcgac 239
           ||||| |||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db     221 ccatcggatgtgcccagatgggattagctagtaggtggggtaacggctcacctaggcgac 280

Qy     240 gatccctagctggtctgagaggatgaccagccacactggaactgagacacggtccagact 299
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     281 gatccctagctggtctgagaggatgaccagccacactggaactgagacacggtttagact 340

Qy     300 cctacgggaggcagcagtgggggaattttgcacaatgggcgcaagcctgatgcagccatgc 359
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     341 cctacgggaggcagcagtgggggaattttgcacaatgggcgcaagcctgatgcagccatgc 400

Qy     360 cgcgtgtatgaagaaggccttcgggttgtaaagtactttcagcggggaggaagggagtaa 419
           ||||||||||||||||||||||||||||||||||||||||||||||||||||  | | 
Db     401 cgcgtgtatgaagaaggccttcgggttgtaaagtactttcagcgggggaggaagggagtaa 460

Qy     420 ggttaataacctcagcgattgacgttacccgcagaagaagcaccggctaactccgtgcca 479
           |||||||  | |     ||||||||||||||||||||||||||||||||||||||||||
Db     461 agttaatacctttgctcattgacgttacccgcagaagaagcaccggctaactccgtgcca 520

Qy     480 gcagccgcggtaatacggagggtgcaagcgttaatcggaattactgggcgtaaagcgcac 539
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     521 gcagccgcggtaatacggagggtgcaagcgttaatcggaattactgggcgtaaagcgcac 580

Qy     540 gcaggcggtctgtcaagtcggatgtgaaatccccgggctcaacctgggaactgcattcga 599
           |||||||| ||| ||||| ||||||||||||||||||||||||||||||||||  ||
Db     581 gcaggcggtttgttaagtcagatgtgaaatccccgggctcaacctgggaactgcatctga 640

Qy     600 aactggcaggctagagtcttgtagaggggggtagaattccaggtgtagcggtgaaatgcg 659
           |||||||  ||| |||||| ||||||||||||||||||||||||||||||||||||||
Db     641 tactggcaagcttgagtctcgtagaggggggtagaattccaggtgtagcggtgaaatgcg 700

Qy     660 tagagatctggaggaataccggtggcgaaggcggccccctggacaaagactgacgctcag 719
           |||||||| |||||||||||||||||||||||||||||||||||||| ||||||||||||
Db     701 tagagatctggaggaataccgg ggcgaaggcggccccctggacgaagactgacgctcag 760

Qy     720 gtgcgaaagcgtggggagcaaacaggattagataccctggtagtccacgccgtaaacgat 779
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     761 gtgcgaaagcgtggggagcaaacaggattagataccctggtagtccacgccgtaaacgat 820

Qy     780 gtcgacttggaggttgtgcccttgaggcgtggcttccggagctaacgcgttaagtcgacc 839
           |||||||||||||||||||||||||||||||||||||||||||||||:|||||||||||
Db     821 gtcgacttggaggttgtgcccttgaggcgtggcttccggagctaacgcgttaagtcgacc 880

Qy     840 gcctggggagtacggccgcaaggttaaaactcaaatgaattgacgggggcccgcacaagc 899
           | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     881 gtctggggagtacggccgcaaggttaaaactcaaatgaattgacgggggcccgcacaagc 940

Qy     900 ggtggagcatgtggtttaattcgatgcaacgcgaagaaccttacctactcttgacatcca 959
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     941 ggtggagcatgtggtttaattcgatgcaacgcgaagaaccttacctggtcttgacatcca 1000

Qy     960 gagaactttccagagatggattggtgccttcgggaactctgagacaggtgctgcatggct 1019
           ||| |.| ||||||||     |||||||||||||||  ||||||||||||||||||||
Db    1001 cggaagttttcagagatgagaatgtgccttcgggaaccgtgagacaggtgctgcatggct 1060

Qy    1020 gtcgtcagctcgtgttgtgaaatgttgggttaagtcccgcaacgagcgcaaccctttatcc 1079
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1061 gtcgtcagctcgtgttgtgaaatgttgggttaagtcccgcaacgagcgcaaccctttatcc 1120

Qy    1080 tttgttgccagcggttaggccgggaactcaaaggagactgccagtgataaactgtgaggaa 1139
           ||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||
Db    1121 tttgttgccagcggtccgccgggaactcaaaggagactgccagtgataaactgtgaggaa 1180

Qy    1140 ggtgaggatgacgtcaagtcatcatggcccttacgacctagggctacacacgtgctacaat 1199
           |||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db    1181 ggtgaggatgacgtcaagtcatcatggcccttacggccagggctacacacgtgctacaat 1240

Qy    1200 ggcgcataccaaagagcgcgacctcgcgagagcaagcggacctcataaagtatgtcgtag 1259
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1241 ggcgcataccaaagagcgcgacctcgcgagagcaagcggacctcataaagtatgtcgtag 1300

Qy    1260 tccggattggagtctgcaactcgactccatgaagtcggaatcgctagtaatcgtagatca 1319
           |||||||||||||||||||.|||||||||||||||||||||||||||||||||||||||
Db    1301 tccggattggagtctgcaactcgactccatgaagtcggaatcgctagtaatcgtagatca 1360

Qy    1320 gaatgctacggtgaatacgttcccgggccttgtacacaccgcccgtcaa 1368
           |||||| |||||||||||| |||||||||||  |||||||||||
Db    1361 gaatgccacggtgaatacgttcccgggccttgcacacaccgcccgtca 1408


RESULT  12
AAN91514
ID   AAN91514 standard; rRNA; 1540 BP.
XX
AC   AAN91514;
XX
DT   02-MAR-1990  (first entry)
XX
DE   Escherichia coli 16S ribosomal RNA.
XX
KW   Probe; Neisseria gonorrhoeae.
XX
OS   Escherichia coli.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    69..99
FT                   /*tag= a
FT   misc_feature    432..488
FT                   /*tag= b
FT   misc_feature    826..858
FT                   /*tag= c
XX
PN   EP337896-A.
XX
PD   18-OCT-1989.
XX
PF   14 APR 1989;   89EP-0401045.
XX
PR   15-APR-1988;   88EP-0400929.
XX
PA   (INNO ) INNOGENETICS NV SA.
XX
PI   Rossau R, Van Heuverswijn;
XX
DR   WPI; 1989-303611/42.
XX
PT   Hybridisation probes for detecting Neisseria strains
PT     some being specific for Neisseria gonorrhoeae
XX
PS   Disclosure; fig. 8; 60pp; English.
XX
CC   N. gonorrhoea rRNA is aligned with this sequence, so specific probe
CC   regions (see features) can be identified
XX
SQ   Sequence 1540 BP; 508 A; 352 C; 487 G; 311 T; 0 other;


  Query Match             85.98%;  Score 1249.2;  DB 10;  Length 1540;
  Best Local Similarity   76.8%;  Pred. No. 0;
  Matches 1050;  Conservative  256;  Mismatches   66;  Indels   27;  Gaps   2;

Qy       1 gacgaacctaaacatctcaa-tcgaacggcagcggaaagtagcttgctactttgccggca 59
           |||||||:|||||:|||| :||| || | | | ||| |||:..: ||||.. .  :
Db      41 ggcaggccaaacacaacggcaagtcgaacggcagcgcaaaagtagcttgctcctttgccgaa 100

Qy      60 agcggcggaacgggtgagtaatgtctgggaaactgcctgatggagggggataactactgga 119
```



RESULT 13
AAS11034
ID AAS11034 standard; DNA; 1487 BP.
XX
AC AAS11034;
XX
DT 24-OCT-2001 (first entry)
XX
DE Shigella dysenteriae 16s ribosomal RNA gene.
XX
KW Antisense; bacterial 16s ribosomal RNA; rRNA; bacterial infection;
KW human; food grain supplement; livestock; poultry; therapeutic; ss.
XX
OS Shigella dysenteriae.
XX
PN WO200142457-A2.
XX
PD 14-JUN-2001.
XX
PF 29-NOV-2000; 2000WO-US42391.
XX
PR 29-NOV-1999; 99US-0168150.
XX
PA (AVIB-) AVI BIOPHARMA INC.
XX
PI Iversen PL;
XX
DR WPI; 2001-457295/49.
XX
PT Antibacterial compound, useful for treating bacterial infections such as
PT livestock and poultry, food supplement, comprises oligonucleotide
PT oligonucleotides complementary to bacterial 16S and 23S rRNA.
XX
PS Disclosure; Page - ; 62pp; English.
XX
CC AAS11021-AAS11034 represent the coding sequences of bacterial 16s
CC ribosomal RNA (rRNA) genes. The sequences were used to design anti-
CC bacterial compounds comprising substantially uncharged antisense
CC oligomers containing 8-40 nucleotide subunits, including a targeting
CC nucleic acid sequence at least 10 nucleotides in length which is
CC complementary to a bacterial 16S or 23S rRNA nucleic acid sequence.
CC The antisense oligomers are used for treating a bacterial infection
CC in a human or a mammalian animal produced by Escherichia coli, Salmonella
CC typhimurium, Pseudomonas aeruginosa, Vibrio cholera, Neisseria
CC gonorrhoea, Helicobacter pylori, Bartonella henselae, Haemophilus
CC influenza, Shigella dysenteriae, Staphylococcus aureus, Mycobacterium
CC tuberculosis, Streptococcus pneumoniae, Treponema pallidum and Chlamydia
CC trachomatis. The antibacterial compound may be used as a food grain
CC supplement in livestock and poultry food composition.
CC Note: The present sequence is not shown in the specification but was
CC been accessed from Genbank using the appropriate accession number given
CC in the specification.
XX
SQ Sequence 1487 BP; 375 A; 341 C; 472 G; 299 T; 0 other;

Query Match 85.7%; Score 1247.6; DB 22; Length 1487;
Best Local Similarity 95.4%; Pred. No. 0;

```
  Matches 1305;  Conservative    0;  Mismatches   61;  Indels    2;  Gaps    2;

Qy       1 ggcaggcctaacacatgcaa-tcgagcggcagcggaaagtagcttgctactttgccaggcg 59
           ||||||||||||||||||||||||| |||| |||| ||| | | ||||| |||||||| ||||| | ||
Db      22 ggcaggcctaacacatgcaagtcgaacggtaacaggaaagcagcttgct-gtttcctgacg 80

Qy      60 agcggcggacgggtgagtaatgtctgggaaactgcctgatggagggggataactactgga 119
           || |||||||||||||||||||||||||||||||| ||||| ||  |||||| ||||||||
Db      81 agtggcggacgggtgagtaatgtctgggaaactgcctgatggagggggataactactgga 140

Qy     120 aacagttgctaataccgcataacgtctcaagacaaaagaggggaccttcgggcctcttg 179
           ||| || ||||||||||||||||| ||||||||||||||||||||||||||||||||| |
Db     141 aacggtagctaataccgcataacgtcgcaagaccaaagagggggaccttcgggcctcttg 200

Qy     180 ccatcagatgtgcccagatgggattagctagtaggtggggtaacggctcacctaggcgac 239
           ||||| |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     201 ccatcggatgtgcccagatgggattagctagtaggtggggtaacggctcacctaggcgac 260

Qy     240 gatccctagctggtctgagaggatgaccagccacactggaactgagacacggtccagact 299
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     261 gatccctagctggtctgagaggatgaccagccacactggaactgagacacggtccagact 320

Qy     300 cctacgggaggcagcagtggggaattgcacaatgggcgcaagcctgatgcagccatgc 359
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     321 cctacgggaggcagcagtggggaattgcacaatgggcgcaagcctgatgcagccatgc 380

Qy     360 cgcgtgtatgaagaaggccttcgggttgtaaagtactttcagcggggaggaagggagtga 419
           ||||||||||||||||||||||||||||||||||||||||||||||||||||     |
Db     381 cgcgtgtatgaagaaggccttcgggttgtaaagtactttcagcggggaggaagggagtaa 440

Qy     420 ggttaataacctcagcgattgacgttacccgcagaagaagcaccggctaactccgtgcca 479
           ||||||| | |     ||||||||||||||||||||||||||||||||||||||||||||
Db     441 agttaatacctttgctcattgacgttacccgcagaagaagcaccggctaactccgtgcca 500

Qy     480 gcagccgcggtaatacggagggtgcaagcgttaatcggaattactgggcgtaaagcgcac 539
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     501 gcagccgcggtaatacggagggtgcaagcgttaatcggaattactgggcgtaaagcgcac 560

Qy     540 gcaggcggtctgtcaagtcggatgtgaaatccccgggctcaacctgggaactgcattcga 599
           |||||||||  ||| ||||| ||||||||||||||||||||||||||||||||  ||
Db     561 gcaggcggtttgttaagtcagatgtgaaatccccgggctcaacctgggaactgcatctga 620

Qy     600 aactggcaggctagagtcttgtagaggggggtagaattccaggtgtagcggtgaaatgcg 659
           ||||||| ||| |||||| ||||||||||||||||||||||||||||||||||||||||||
Db     621 tactggcaagcttgagtctcgtagaggggggtagaattccaggtgtagcggtgaaatgcg 680

Qy     660 tagagatctggaggaataccggtggcgaaggcggccccctgtacaaaagactgacgctcag 719
           ||||||||||||||||||||||||||||||||||||||||||| || ||||||||||||
Db     681 tagagatctggaggaataccggtggcgaaggcggccccctgtacaaaaactgacgctcag 740

Qy     720 gtgcgaaagcgtggggagcaaacaggattagataccctggtagtccacgccgtaaacgat 779
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     741 gtgcgaaagcgtggggagcaaacaggattagataccctggtagtccacgccgtaaacgat 800

Qy     780 gtcgacttggaggttgtgcccttgaggcgtggcttccggagctaacgcgttaagtcgacc 839
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     801 gtcgacttggaggttgtgcccttgaggcgtggcttccggagctaacgcgttaagtcgacc 860

Qy     840 gcctggggagtacggccgcaaggttaaaactcaaatgaattgacgggggcccgcacaagc 899
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     861 gcctggggagtacggccgcaaggttaaaactcaaatgaattgacgggggcccgcacaagc 920

Qy     900 ggtggagcatgtggtttaattcgatgcaacgcgaagaaccttacctactcttgacatcca 959
           |||||||||||||||||||||||||||||||||||||||||  ||||||||||||
Db     921 ggtggagcatgtggtttaattcgatgcaacgcgaagaaccttacctggtcttgacatcca 980

Qy     960 gagaactttccagagatggattggtgccttcgggaactctgagacaggtgctgcatggct 1019
           ||||| ||  ||||||    ||||||||||||||| ||||||||||||||||||||||||
Db     981 cagaacttgtagagatacgaggtggtgccttcgggaactgtgagacaggtgctgcatggct 1040

Qy    1020 gtcgtcagctcgtgttgtgaaatgttgggttaagtcccgcaacgagcgcaaccctt-tcc 1078
           ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||| |||
Db    1041 gtcgtcagctcgtgttgtgaaatgttgggttaagtcccgcaacgagcgcaaccctttatcc 1100

Qy    1079 ttgttgccagcggttaggccgggaactcaaaggagactgccagtgataaactggaggaag 1138
           ||||||||||||||  ||||||||||||||||||||||||||  ||| ||||||||||||
Db    1101 ttgttgccagcggtccggccgggaactcaaaggagactgccagtgataaactgtaggaag 1160

Qy    1139 gtggggatgacgtcaagtcatcatggcccttacgagtagggctacacacgtgctacaat 1198
           ||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db    1161 gtggggatgacgtcaagtcatcatggcccttacggccagggctacacacgtgctacaat 1220

Qy    1199 ggccgtacaaagagaagcgacctcgcgagagcaagcggacctcataaagtcggtcgtagtc 1258
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1221 ggccgtacaaagagaagcgacctcgcgagagcaagcggacctcataaagtcggtcgtagtc 1280

Qy    1259 tccggattggagtctgcaactcgactccatgaagtcggaatcgctagtaatcgcgaatca 1318
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1281 tccggattggagtctgcaactcgactccatgaagtcggaatcgctagtaatcgcgaatca 1340

Qy    1319 gaatgtcacggtgaatacgttcccgggccttgtacacaccgcccgtcaca 1367
           |||||  |||||||||||||||||||||||||||||||||||||||||||
Db    1341 gaatgtcacggtgaatacgttcccgggccttgtacacaccgcccgtcaca 1389


RESULT  14
AAX24986/c
ID   AAX24986 standard; DNA; 5341 BP.
XX
AC   AAX24986;
XX
DT   05-JUL-1999  (first entry)
XX
DE   E. coli MG1655 rrnD operon (16S-spacer-23S-spacer-5S).
XX
KW   Applications; Diagnostic; species-specific; eliminator; marker; RFLP;
KW   restriction fragment length polymorphism; bacterium; fungus;
KW   pathogen; rrnD operon; 16S RNA gene; 23S RNA gene; 5S.
XX
OS   Escherichia coli.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    complement (1..121)
FT                   /*tag= a
FT                   /label= 5S
FT   misc_feature    complement (449..3362)
FT                   /*tag= b
FT                   /label= 23S
FT   misc_feature    complement (3800..4341)
FT                   /*tag= c
FT                   /label= 16S
FT
XX
PN   WO9905425-A1.
XX
PD   04-FEB-1999.
XX
PF   24-JUL-1998;    98WO-US15464.
XX
PR   25-JUL-1997;    97US-0053097.
XX
PA   (UYBO-) UNIV BOSTON.
XX
PI   Goldstein RN;
XX
DR   WPI; 1999-142969/12.
XX
PT   Determining species of bacteria and fungi - useful for
PT   distinguishing between bacterial/fungal species, and for determining
PT   the identity of bacterial/fungal pathogens in biological samples
XX
PS   Disclosure; Fig 7 (39/67 42/67); 143pp; English.
XX
CC   This is the DNA sequence of the Escherichia coli strain MG1655
CC   rrnD operon (16S-spacer-23S-spacer-5S).  Restriction sites for
CC   enzymes cutting the operon 5 times or less have been determined.
CC   E. coli rrnA-rrnH operon sequences are provided (see AAX24983-8).
```

```
CC   Methods and compositions are described for determining the species
CC   of an unknown bacterium or fungus in a sample.  The method involves
CC   isolating and digesting bacterial (or fungal) DNA encoding 16S and
CC   23S rRNA from a sample with restriction enzymes, detecting the
CC   products, and comparing them to signature bands from a number of
CC   bacteria.  The method generates a species conserved set of RFLP
CC   bands, unique for each species.  These species-conserved sets
CC   represent precise markers appropriate for inter-species
CC   discriminatory purposes (i.e. to determine the species of a given,
CC   unknown isolate e.g. in a clinical specimen).  In contrast to
CC   conventional ribotyping, the present invention utilises the
CC   ribosomal operon sequences which vary less than 3% (and more
CC   preferably less than 2%) within a species and vary between species.
CC   The method is useful for medical, food, agricultural and
CC   environmental testing.  It does not require sequencing of nucleic
CC   acid from biological samples.
XX
SQ   Sequence 5341 BP; 1117 A; 1647 C; 1188 G; 1389 T; 0 other;

  Query Match             85.6%;  Score 1245.2;  DB 23;  Length 5341;
  Best Local Similarity   94.8%;  Pred. No. 0.;
  Matches 1297;  Conservative    0;  Mismatches   70;  Indels    1;  Gaps    1;

Qy       . ggcaggcctaacacatgcaa-tcgagcggcagcggaaagtagcttgctacttt-gccggcg 59
           |||||||||||||||||||| |||| ||| | | | ||  |||||||  || || | ||
Db    5301 GGCAGGCCTAACACATGCAAGTCGAACGGTAACAGGAAACAG-TTGCTGTTTTCGGCTGACG 5242

Qy      60 agcggcggacgggtgagtaatgtctgggaaactacctgatggttgtgggatcactactgga 119
           || |||||||||||||||||||||||||||||||||||| ||||| ||  |||||| ||||||||
Db    5241 AGTGGCGGACGGGTGAGTAATGTCTGGGAAACTGCCTGATGGAGGGGGATAACTACTGGA 5182

Qy     120 aacagttgctaataccgcataacgtctcaagaccaaagaggggacctt-cgggcctcttc 179
           ||| || |||||||||||||||||| |||||||||||||||||||| |||||||||||
Db    5181 AACGGTAGCTAATACCGCATAACGTCGCAAGACCAAAGAGGGGGACCT-CGGGCCTCTTG 5122

Qy     180 ccatcagatgtgcccagatgggattagctagtaggtggggtaacggctcacctaggcgac 239
           ||||| |||||||||||||||||||||| || ||||||||||||||||||| |||||||
Db    5121 CCATCGGATGTGCCCAGATGGGATTAGCTTGTTGGTGGGGTAACGGCTCACCAAGGCGAC 5062

Qy     240 gatccctagctggtctgagaggatgaccagccacactggaactgagacacggtccagact 299
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5061 GATCCCTAGCTGGTCTGAGAGGATGACCAGCCACACTGGAACTGAGACACGGTCCAGACT 5002

Qy     300 cctacgggaggcagcagtgggcaatattgcacaatgggcgcaagcctgatgcagccatgc 359
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    5001 CCTACGGGAGGCAGCAGTGGGGAATATTGCACAATGGGCGCAAGCCTGATGCAGCCATGC 4942

Qy     360 cgcgtgtatgaagaaggccttcgggttgtaaagtactttcagcggggaggaagcgcttga 419
           |||||||||||||||||||||||||||||||||||||||||||||||||||   | |
Db    4941 CGCGTGTATGAAGAAGGCCTTCGGGTTGTAAAGTACTTTCAGCGGGGAGGAAGGGAGTAA 4882

Qy     420 ggttaataacctcagcgattgacgttacccgcagaagaagcaccggctaactccgtgcca 479
           |||||||  | |     ||||||||||||||||||||||||||||||||||||||||||
Db    4881 AGTTAATACCTTTGCTCATTGACGTTACCCGCAGAAGAAGCACCGGCTAACTCCGTGCCA 4822

Qy     480 gcagccgcggtaatacggagcgtgcaagcgttaatcggaattactgggcgtaaagcgcac 539
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4821 GCAGCCGCGGTAATACGGAGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAAGCGCAC 4762

Qy     540 gcaggcggtctgtcaagtcggatgtgaaatccccgggctcaacctgggaactgcattcga 599
           ||||||||| ||| ||||| ||||||||||||||||||||||||||||||||  ||
Db    4761 GCAGGCGGTTTGTTAAGTCAGATGTGAAATCCCCGGGCTCAACCTGGGAACTGCATCTGA 4702

Qy     600 aactggcaggctagagtcttgtagaggggggtagaattccaggtgtagcggtgaaatgcg 659
           |||||| ||| |||||| ||||||||||||||||||||||||||||| ||||||||||
Db    4701 TACTGGCAAGCTTGAGTCTCGTAGAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCG 4642

Qy     660 tagagatctggaggaataccggtggcgaaggcggccccctggacgaagactgacgctcag 719
           |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    4641 TAGAGATCTGGAGGAATACCGGTGGCGAAGGCGGCCCCCTGGACGAAGACTGACGCTCAG 4582

Qy     720 gtgcgaaagcgtggggagcaaacaggattagataccctggtagtccacgccgtaaacgat 779
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4581 GTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAACGAT 4522

Qy     780 gtcgacttggaggttgtgcccttgaggcgtggcttccggagctaacgcgttaagtcgacc 839
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4521 GTCGACTTGGAGGTTGTGCCCTTGAGGCGTGGCTTCCGGAGCTAACGCGTTAAGTCGACC 4462

Qy     840 gcctggggagtacggccgcaaggttaaaactcaaatgaattgacgggggcccgcacaagc 899
           |||||||||||||||||||| |||||||||||||||||||||||||||||||||||||||
Db    4461 GCCTGGGGAGTACGGCCGCAAGGTTAAAACTCAAATGAATTGACGGGGGCCCGCACAAGC 4402

Qy     900 ggtggagcatgtggtttaattcgatgcaacgcgaagaaccttacctactcttgacatcca 959
           |||||||||||||||||||||||||||||||||||||||||||||||||||  ||||||||
Db    4401 GGTGGAGCATGTGGTTTAATTCGATGCAACGCGAAGAACCTTACCTGGTCTTGACATCCA 4342

Qy     960 gaagaactttccagagatggattggtgccttcgggaactctgagacaggtgctgcatggct 1019
           ||| ||| |||||||     ||||||||||||||  ||||||||||||||||||||||||
Db    4341 CGGAAGTTTTCAGAGATGAGAATGTGCCTTCGGGAACGCCTGACAGGTGCTGCATGGCT 4282

Qy    1020 gtcgtcagctcgtgttgtgaaatgttgggttaagtcccgcaacgagcgcaacccttatcc 1079
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4281 GTCGTCAGCTCGTGTTGTGAAATGTTGGGTTAAGTCCCGCAACGAGCGCAACCCTTATCC 4222

Qy    1080 tttgttgccagcgattaggccgggaactcaaaggagactgccggtgataaaccggaggaag 1139
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4221 TTTGTTGCCAGCGGTTAGGCCGGGAACTCAAAGGAGACTGCCGGTGATAAACCGGAGGAA 4162

Qy    1140 ggtggggatgacgtcaagtcatcatggcccttacgagtagggctacacacgtgctacaat 1199
           |||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    4161 GGTGGGGATGACGTCAAGTCATCATGGCCCTTACGACCAGGGCTACACACGTGCTACAAT 4102

Qy    1200 ggcgcatacaaagagaagcgacctcgcgagagcaagcggacctcataaagtgcgtcgtag 1259
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4101 GGCGCATACAAAGAGAAGCGACCTCGCGAGAGCAAGCGGACCTCATAAAGTGCGTCGTAG 4042

Qy    1260 tccggatttggagtctgcaactcgactccatgaagtcggaatcgctagtaatcgtggatca 1319
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    4041 TCCGGATTGGAGTCTGCAACTCGACTCCATGAAGTCGGAATCGCTAGTAATCGTGGATCA 3982

Qy    1320 gaatgctacggtgaatacgttcccgggccttgtacacaccgcccgtcaca 1369
           |||||| |||||||||||||||||||||||| |||||||||||||| ||
Db    3981 GAATGCCACGGTGAATACGTTCCCGGGCCGTTGTACACACCGCCCGTCA 3934


RESULT  15
AAX24985
ID   AAX24985 standard; DNA; 5013 BP.
XX
AC   AAX24985;
XX
DT   05-JUL-1999  (first entry)
XX
DE   E. coli MG1655 rrnC operon (16S-spacer-23S-spacer-5S).
XX
KW   Speciation, ribotyping, species discrimination, marker, RFLP;
KW   restriction fragment length polymorphism; bacterium; fungus;
KW   pathogen; rrnC operon; 16S RNA gene; 23S RNA gene; 5S.
XX
OS   Escherichia coli.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1..1541
FT                   /*tag= a
FT                   /label= 16S
FT   misc_feature    1896..4801
FT                   /*tag= b
FT                   /label= 23S
FT   misc_feature    4894..5013
FT                   /*tag= c
FT                   /label= 5S
XX
PN   WO9905325-A1.
XX
```

PD 04-FEB-1999.
XX
PF 24-JUL-1998; 98WO-US15464.
XX
PR 25-JUL-1997; 97US-0053097.
XX
PA (UYBC-) UNIV BOSTON.
XX
PI Goldstein RN;
XX
DR WPI; 1999-142969/12.
XX
PT Determining species of bacteria and fungi - useful for
PT distinguishing between bacterial/fungal species, and for determining
PT the identity of bacterial/fungal pathogens in biological samples
XX
PS Disclosure; Fig 7 (32/67-35/67); 133pp; English.
XX
CC This is the DNA sequence of the Escherichia coli strain MG1655
CC rrnC operon (16S-spacer-23S-spacer-5S). Restriction sites for
CC enzymes cutting the operon 5 times or less have been determined.
CC E. coli rrnA-rrnH operon sequences are provided (see AAX24993-99).
CC Methods and compositions are described for determining the species
CC of an unknown bacterium or fungus in a sample. The method involves
CC isolating and digesting bacterial (or fungal) DNA encoding 16S and
CC 23S rRNA from a sample with restriction enzymes, detecting the
CC products, and comparing them to signature bands from a number of
CC bacteria. The method generates a species conserved set of RFLP
CC bands, unique for each species. These species-conserved sets
CC represent precise markers appropriate for inter species
CC discriminatory purposes (i.e. to determine the species of a given,
CC unknown isolate e.g. in a clinical specimen). In contrast to
CC conventional ribotyping, the present invention utilises the
CC ribosomal operon sequences which vary less than 3% (and more
CC preferably less than 2%) within a species and vary between species.
CC The method is useful for medical, food, agricultural and
CC environmental testing. It does not require sequencing of nucleic
CC acid from biological samples.
XX
SQ Sequence 5015 BP; 1310 A; 1151 C; 1533 G; 1059 T; 0 other;

Query Match 85.4%; Score 1242.8; DB 20; Length 5015;
Best Local Similarity 95.2%; Pred. No. 0;
Matches 1302; Conservative 0; Mismatches 64; Indels 2; Gaps 2;